

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:30 ; Search time 70 Seconds
(without alignments)
5011.310 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSRLGVLLSLPVLQLAT.....PAYPTESCHLSLVAFVPCVL 907

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : A_Geneseq_16Dec04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	2 AAW93889	AAW93889 Human HG3
2	4702	100.0	907	3 AAY90682	Aay90682 Human G P
3	4702	100.0	907	6 ABP81968	Abp81968 Human G P
4	4702	100.0	907	6 ABO06467	AbO06467 Human G-P
5	4702	100.0	907	7 ADC22783	Adc22783 Human G P
6	4702	100.0	907	7 ADE59150	AdE59150 Human Pro
7	4702	100.0	907	7 ADE59153	AdE59153 Human Pro
8	4702	100.0	907	7 ADG42628	AdG42628 Human G P
9	4702	100.0	907	7 ADH14256	AdH14256 Human HG3
10	4702	100.0	907	7 ADN40013	Adn40013 Cancer/an
11	4702	100.0	907	7 ADN39531	Adn39531 Cancer/an
12	4702	100.0	907	7 ADN39628	Adn39628 Cancer/an
13	4702	100.0	907	8 AD029408	Ado29408 Human GPC
14	4702	100.0	907	8 ADQ80369	Adq80369 G protein
15	4702	100.0	907	8 ADG67868	Adg67868 Human HG3
16	4701	100.0	1145	7 ADF70480	Adf70480 Orphan re
17	4696	99.9	907	3 AAY90687	Aay90687 Human mut
18	4696	99.9	907	7 ADC22797	Adc22797 Human G P
19	4696	99.9	907	7 ADH14270	Adh14270 Mutated h
20	4691	99.8	907	7 ADG42629	Adg42629 Human G P
21	4556	96.9	883	7 ADB80464	Adb80464 Ovarian c
22	4556	96.9	883	7 ADN40012	Adn40012 Cancer/an
23	4556	96.9	883	7 ADN39166	Adn39166 Cancer/an
24	4048	86.1	907	8 ADG75449	Adg75449 Mouse orp
25	4048	86.1	907	8 AD029409	Ado29409 Mouse GPC

ALIGNMENTS

RESULT 1

AAW93889

ID AAW93889 standard; protein; 907 AA.

XX AC AAW93889;

XX DT 25-JUN-1999 (first entry)

XX DE Human HG38 protein.

XX KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;

KW endocrine system; skeletal muscle; spinal cord; placenta; development;

KW receptor activity modulator.

XX OS Homo sapiens.

XX PN WO9915660-A1.

XX PD 01-APR-1999.

XX PF 24-SEP-1998; 98WO-US019979.

XX PR 24-SEP-1997; 97US-0059863P.

XX XX (MERI) MERCK & CO INC.

XX LIU Q, Bailey WJ, McDonald TP;

XX WPI; 1999-254711/21.

XX N-PSDB; AAX23980.

XX Human G-protein coupled glycoprotein hormone receptor HG38.

XX Claim 1a; Fig 2; 74pp; English.

This invention describes a novel human G-protein coupled glycoprotein hormone receptor, HG38. Glycoprotein hormone receptors are important in the endocrine system and HG38 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the HG38 receptor, as well as for studying the ability of a variety of compounds to act as modulators of HG38 receptor activity

SQ Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 2; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPCHCEPDGRMLLRVDCSDLGSEL 60
Db	1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPCHCEPDGRMLLRVDCSDLGSEL 60
Qy	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRLAGNALTYPKGAFTGLSKLVLM 120
Db	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRLAGNALTYPKGAFTGLSKLVLM 120
Qy	121 LQNNQLRHVPTEALQNLRSLSRLDANHISYVPPSPCSGLHSRLHRLWLDNALTETIPVQ 180
Db	121 LQNNQLRHVPTEALQNLRSLSRLDANHISYVPPSPCSGLHSRLHRLWLDNALTETIPVQ 180
Qy	181 AFRSLSALQMTALNKIHHPDYAFGNLSLVVLHNNRIHSLGKCFGLHSLETLTD 240
Db	181 AFRSLSALQMTALNKIHHPDYAFGNLSLVVLHNNRIHSLGKCFGLHSLETLTD 240
Qy	241 LNYNNLDEFFTAIRTLNKLKELGFHNSNNIRSIPEKAFVGNPSLTIHFYDNPQFVGSA 300
Db	241 LNYNNLDEFFTAIRTLNKLKELGFHNSNNIRSIPEKAFVGNPSLTIHFYDNPQFVGSA 300
Qy	301 FOHLPELRTLNGASQITTEPDDLTGTANLESLLTGAQISSLPQTVNCQNPNIQVLDLS 360
Db	301 FOHLPELRTLNGASQITTEPDDLTGTANLESLLTGAQISSLPQTVNCQNPNIQVLDLS 360
Qy	361 YNLLEDLPFSVFCOKLQIDIRHNEIYEIKVDYTFQQLLSLSRLANWKIAIHPNAPST 420
Db	361 YNLLEDLPFSVFCOKLQIDIRHNEIYEIKVDYTFQQLLSLSRLANWKIAIHPNAPST 420
Qy	421 LPSLIKLDLSNLLSSFFITGLHGLTHLKGHALQSLISSENFPELPKVIEMPVAYQCC 480
Db	421 LPSLIKLDLSNLLSSFFITGLHGLTHLKGHALQSLISSENFPELPKVIEMPVAYQCC 480
Qy	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Qy	541 CSPSPGPKCEHLLDGLHIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db	541 CSPSPGPKCEHLLDGLHIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy	601 AVNMLTGVSAAVLGADVAFSGSFARHAGWENGCHVIGFLSIFASESSVFLITLAAL 660
Db	601 AVNMLTGVSAAVLGADVAFSGSFARHAGWENGCHVIGFLSIFASESSVFLITLAAL 660
Qy	661 ERGFSVKYSAPFETKAPFSSLKVIILLCALLATMAAVPLLGSGKYGASPICLPLPFGEP 720
Db	661 ERGFSVKYSAPFETKAPFSSLKVIILLCALLATMAAVPLLGSGKYGASPICLPLPFGEP 720
Qy	721 STMGVWVALIILNSICFLMMTIAVTKLYCNLDKGDLENIWDCSMVKHTALLFTNCILNC 780
Db	721 STMGVWVALIILNSICFLMMTIAVTKLYCNLDKGDLENIWDCSMVKHTALLFTNCILNC 780
Qy	781 PVAFLSFLSINLFTFISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db	781 PVAFLSFLSINLFTFISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy	901 VAFVPCL 907
Db	901 VAFVPCL 907

RESULT 2
ID AA90682
XX standard; protein; 907 AA.
AC AA90682;
XX

DT 21-AUG-2000 (first entry)
XX Human G protein-coupled receptor HG38.
XX
XX G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist.
XX
XX Homo sapiens.
XX WO200022129-A1.
XX 20-APR-2000.
XX 12-OCT-1999; 99WO-US023938.
XX 13-OCT-1998; 98US-00170496.
XX (AREN-) ARENA PHARM INC.
XX Behan DP, Chalmers DT, Liaw CW;
XX WPI; 2000-329165/28.
XX N-ESDB; AAA30770.
XX Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
PT agents.
XX
XX Example 1; Page 317-320; 341pp; English.
XX
XX The invention relates to constitutively active, non-endogenous versions
CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
CC and AAA30775-A30779). The mutant proteins of the invention contain a
CC mutation in a portion of the protein comprising intracellular loop 3
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC is substituted for an endogenous residue in IC3 at a position 16 amino
CC acids N-terminal of an endogenous proline in TM6 to form a sequence X-
CC (AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or
CC Ala, and is preferably Lys. When the endogenous residue at this position
CC is Lys, this residue is replaced by His, Arg or preferably Ala. The 15
CC amino acid stretch between the substituted amino acid and the Pro may be
CC endogenous, non-endogenous, or a mixture of endogenous and non-endogenous
CC residues. The constitutively active GPCRs are useful for identifying
CC antagonists, agonists and partial agonists for use as pharmaceutical
CC agents. The mutant proteins are also useful in research settings for
CC elucidating the roles of the receptors in normal and diseased conditions.
CC Antagonists for a particular GPCR are useful for treating diseases and
CC disorders associated with that receptor. Because the novel mutant GPCRs
CC are constitutively active, they can be used directly for screening of
CC compounds without the need for endogenous ligands. The present sequence
CC represents a human wild-type GPCR referred to in an exemplification of
CC the invention
XX
XX Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 3; Length 907;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPCHCEPDGRMLLRVDCSDLGSEL 60
Db	1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPCHCEPDGRMLLRVDCSDLGSEL 60
Qy	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRLAGNALTYPKGAFTGLSKLVLM 120
Db	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRLAGNALTYPKGAFTGLSKLVLM 120
Qy	121 LQNNQLRHVPTEALQNLRSLSRLDANHISYVPPSPCSGLHSRLHRLWLDNALTETIPVQ 180
Db	121 LQNNQLRHVPTEALQNLRSLSRLDANHISYVPPSPCSGLHSRLHRLWLDNALTETIPVQ 180

QY 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVLHNNRHSLSGKCFDGLHSLETLD 240
Db 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVLHNNRHSLSGKCFDGLHSLETLD 240
QY 241 LANNLNDEPPTAIRTLNLSKELGFSNNIRSIPEKAFVGNPSLIITHFYDNPITQFVGRSA 300
Db 241 LANNLNDEPPTAIRTLNLSKELGFSNNIRSIPEKAFVGNPSLIITHFYDNPITQFVGRSA 300
QY 301 FOHLPELRITLNGASQITTEPDLTGTANLESILITGAQISSLPQVNCNLPNQLVDLS 360
Db 301 FOHLPELRITLNGASQITTEPDLTGTANLESILITGAQISSLPQVNCNLPNQLVDLS 360
QY 361 YNLLEDLPFSVCOKLOKIDLRHNEYIYKVTFOQLLSIRSLNLANWKIAIHNNAFT 420
Db 361 YNLLEDLPFSVCOKLOKIDLRHNEYIYKVTFOQLLSIRSLNLANWKIAIHNNAFT 420
QY 421 LPSLIKDLSSNLSFPITGLHGLTHLKLGTGNHALQSLSISSENPELVIMPEYAYQCC 480
Db 421 LPSLIKDLSSNLSFPITGLHGLTHLKLGTGNHALQSLSISSENPELVIMPEYAYQCC 480
QY 481 AFGVCENAYKISQNMWNGDSSMDLHKKDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNMWNGDSSMDLHKKDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
QY 541 CSPSPGPPKCEHLLDGNLIRIGWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPPKCEHLLDGNLIRIGWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLADVATFGSPARHGAWENGCHVIGFLSIPASSSSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLADVATFGSPARHGAWENGCHVIGFLSIPASSSSVFLTLAAL 660
QY 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATWAAVPLLGSKYKASPLCLPFPGBP 720
Db 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATWAAVPLLGSKYKASPLCLPFPGBP 720
QY 721 STMGYVALILNLSICFLMWTATYKLYCNLDKGLNWDSCMWKHIALLFTNCILNC 780
Db 721 STMGYVALILNLSICFLMWTATYKLYCNLDKGLNWDSCMWKHIALLFTNCILNC 780
QY 781 PVAFISFSSLINLTFTSPVIFKILLVVVPLPACLNPLIYILFNPHFKEDLVSLRKQTV 840
Db 781 PVAFISFSSLINLTFTSPVIFKILLVVVPLPACLNPLIYILFNPHFKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDVQKSCDSTQALVFTSSSITDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVQKSCDSTQALVFTSSSITDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 3
ID ABP81968 standard; protein; 907 AA.
XX AC ABP81968;
XX DT 04-MAR-2003 (first entry)
XX DE Human G protein-coupled receptor GPR49 protein SEQ ID NO:422.
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

ulcer.
XX Homo sapiens.
XX WO200261087-A2.
XX 08-AUG-2002.
PF 19-DEC-2001; 2001WO-US050107.
XX 19-DEC-2000; 2000US-0257144P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX Burner GC, Roush CL, Brown JP;
XX WPI: 2003-046718/04.
XX N-PSDB; AB242816.
102 e 7
not-species fr

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.
Disclosure; Fig 1; 523pp; English.
The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (1) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial lung, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 6; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSELGVLISLPVLLQATGSSPRSGVLLRGCPHCEPDGRMLLRVDCSDGLSSEL 60
Db 1 MDTSELGVLISLPVLLQATGSSPRSGVLLRGCPHCEPDGRMLLRVDCSDGLSSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNLPISLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNLPISLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSIVPPSPCSFGLSHRLMDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSIVPPSPCSFGLSHRLMDDNALTEIPVQ 180
QY 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVLHNNRHSLSGKCFDGLHSLETLD 240
Db 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVLHNNRHSLSGKCFDGLHSLETLD 240

QY 241 LNNLDEFFTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
DB 241 LNNLDEFFTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESLLTGAQISSLPQTVCNQLPNIQVLDLS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESLLTGAQISSLPQTVCNQLPNIQVLDLS 360
QY 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDTTFOQLLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDTTFOQLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLTKLDSNLSSFFITGLHGTHTLKGNAHQSLISSENPPELKVEMPVAYOCC 480
DB 421 LPSLTKLDSNLSSFFITGLHGTHTLKGNAHQSLISSENPPELKVEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDLHKKADAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKKADAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKCEHLLDGLHIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPKCEHLLDGLHIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSAAVLGADVFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSAAVLGADVFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGSVKYSKAFETKAPFSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720
DB 661 ERGSVKYSKAFETKAPFSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720
QY 721 STMGYVALIILNSLCFLMTIATKLYCNLDKGDLENWCSWKHTALLFTNCILNC 780
DB 721 STMGYVALIILNSLCFLMTIATKLYCNLDKGDLENWCSWKHTALLFTNCILNC 780
QY 781 PVALFSSSLNLTPISEVIFKILLVVPACLNPLLYLFNPHKEDLVSLRKQTYV 840
DB 781 PVALFSSSLNLTPISEVIFKILLVVPACLNPLLYLFNPHKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQCDSTQALVTTSSITVDLPSSVPSPAYPVTTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQCDSTQALVTTSSITVDLPSSVPSPAYPVTTESCHLSS 900
QY 901 VAFVPCCL 907
DB 901 VAFVPCCL 907

RESULT 4
ABO06467
ID ABO06467 standard; protein; 907 AA.
XX AC ABO06467;
XX DT 13-AUG-2003 (first entry)
XX DE Human G-protein coupled receptor HG38.
XX KW G-protein coupled receptor; GPCR; caudate nucleus related disorder;
KW neoplastic disorder; leukaemia; breast cancer; immune disorder; AIDS;
KW rheumatoid arthritis; neuronal disease; Alzheimer's disease; receptor;
KW Parkinson's disease; respiratory disorder; bronchopulmonary disease;
KW pancreatic disease; ovarian disorder; ovarian carcinoma; colon disease;
KW Meig's syndrome; breast disease; thalamus related disorder; gene therapy;
KW amygdala related disorder; corpus callosum related disorder;
XX KW hippocampus related disorder.
OS Homo sapiens.
XX FN US2003027323-A1.
XX

PD 06-FEB-2003.
XX 26-SEP-2001; 2001US-00965536.
XX 27-SEP-2000; 2000US-0235832P.
PR 16-JAN-2001; 2001US-0261781P.
PR 19-JUL-2001; 2001US-0306605P.
PR 03-AUG-2001; 2001US-0310436P.
XX (FEDE/) FEDER J N.
PA (MINT/) MINTIER G.
PA (RAMA/) RAMANATHAN C S.
PA (HAWK/) HAWKEN D R.
XX Feder JN, Minter G, Ramanathan CS, Hawken DR;
XX WPI; 2003-466147/44.
DR New G-protein coupled receptor polypeptides, designated as HGPRMY5,
XX useful for preventing, treating or ameliorating a medical condition
XX related to the colon, breast, ovaries or immune system.
XX Disclosure; Fig 8; 93pp; English.
XX The invention relates to an isolated HGPRMY5 polypeptide. The
XX polypeptides, polynucleotides and methods are useful for preventing,
XX treating or ameliorating a medical condition such as a neoplastic
XX disorder e.g. leukaemia and breast cancer; immune disorder e.g. AIDS and
XX rheumatoid arthritis; neuronal disorder e.g. Alzheimer's disease and
XX Parkinson's disease; respiratory disorder e.g. bronchopulmonary disease
XX and pancreatic disease; ovarian disorder e.g. ovarian carcinoma and
XX Meig's syndrome; colon disease; breast disease; thalamus related disorder;
XX amygdala related disorder; corpus callosum related disorder; caudate
XX nucleus related disorder; hippocampus related disorder by administering
XX the GPCR polypeptide or its homologue. The present sequence represents
XX the amino acid sequence of a G-protein coupled receptor used to show
XX homology with the human G-protein coupled receptor, HGCRBMY
XX Sequence 907 AA;
XX
XX Query Match 100.0%; Score 4702; DB 6; Length 907;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
DB 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRFLERLRLAGNALTYPKGAFGLYSKVLIM 120
DB 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRFLERLRLAGNALTYPKGAFGLYSKVLIM 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRDANDHISYVPPSCFSGHLSRLHLDNALTETIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRDANDHISYVPPSCFSGHLSRLHLDNALTETIPVQ 180
QY 181 AFRSLSAQMTALNKLTHIIPDYAFGNLSLVVLHLHNNRHISLGKCKPGLHSLETLD 240
DB 181 AFRSLSAQMTALNKLTHIIPDYAFGNLSLVVLHLHNNRHISLGKCKPGLHSLETLD 240
QY 241 LNNYNLDEFFTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
DB 241 LNNYNLDEFFTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESLLTGAQISSLPQTVCNQLPNIQVLDLS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESLLTGAQISSLPQTVCNQLPNIQVLDLS 360
QY 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDTTFOQLLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDTTFOQLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLTKLDSNLSSFFITGLHGTHTLKGNAHQSLISSENPPELKVEMPVAYOCC 480

check

Db 421 LPSLKLDSNLLSPFITGLGTHLKLGNHALQSLISENPELKIEMPIYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGLWIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGLWIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSSAVLAGVDAFTGSPARHGAWNGVGVCHVIGFLSPASESSVFLTLTAAAL 660
Db 601 AVNMLTGVSSAVLAGVDAFTGSPARHGAWNGVGVCHVIGFLSPASESSVFLTLTAAAL 660
QY 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLTMAAVPLLGSGKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLTMAAVPLLGSGKYGASPLCLPLPFGEP 720
QY 721 STMGWVVALIILNSLCFLMTTAYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGWVVALIILNSLCFLMTTAYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFSLSPSSLINLTETISPEVIKFIILLVVVPLPACLNPLLYILFNPHEKEDVSLRKQTVV 840
Db 781 PVAFSLSPSSLINLTETISPEVIKFIILLVVVPLPACLNPLLYILFNPHEKEDVSLRKQTVV 840
QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTECHLSS 900
Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTECHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 5
ADC22783
ID ADC22783 standard; protein; 907 AA.
XX
AC ADC22783;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human G protein-coupled receptor (GPCR) polypeptide #39.
XX
KW Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6;
KW intracellular-3 region; IC3; receptor.
XX
OS Homo sapiens.
XX
PN US6555339-B1.
XX
PD 29-APR-2003.
XX
PF 13-OCT-1998; 98US-00170498.
XX
PR 14-APR-1997; 97US-00839449.
PR 14-APR-1998; 98US-00060188.
PR 26-JUN-1998; 98US-0090783P.
PR 07-AUG-1998; 98US-0095677P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Liaw CW, Behan DP, Chalmers DT;
XX
XX WPI; 2003-742861/70.
DR N-PSDB; ADC22782.
DR
XX
PT Creating a constitutively active version of an endogenous human G protein
PT coupled receptor (GPCR) comprises substituting a specific amino acid in
PT the transmembrane-6 region with a different amino acid, and testing for
PT constitutive activity.

XX Example 1; SEQ ID NO 264; 221pp; English.
PS The invention relates to a method for treating a non-endogenous,
CC constitutively active version of an endogenous human G protein-coupled
CC receptor (GPCR) that has a transmembrane-6 (TM6) region and an
CC intracellular-3 (IC3) region, by substituting a specific amino acid in
CC the TM6 region with a different amino acid, and testing for constitutive
CC activity. The method is useful for creating a constitutively active
CC version of an endogenous human GPCR that comprises a transmembrane 6
CC region and an intracellular loop 3 region. The altered human GPCR
CC polypeptides are useful for screening test compounds for identification
CC of inverse agonists or partial agonists of GPCR polypeptides, which may
CC have therapeutic uses. The altered GPCRs may also be used in vivo or in
CC vitro in biological research. A nucleic acid encoding the altered GPCR
CC may be used to create a transgenic animal expressing the altered GPCR.
CC The method allows screening for compounds that modulate the activity of a
CC human G protein-coupled receptor without the need for provision of a
CC ligand for the receptor. This is particularly useful in allowing
CC screening of compounds against orphan receptors for which no ligand is
CC currently known. This sequence represents a human GPCR polypeptide of the
CC invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRLGVLISLPVLLQATGSSPSRSGVLGRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRLGVLISLPVLLQATGSSPSRSGVLGRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLLEELRAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLLEELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRLDANHI SYVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSQSLRLDANHI SYVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
QY 181 AFRSLSALQMTALNKLHHPDYAFGNLSLVVLHNNRTHSLGKCFDGLHSLTLD 240
Db 181 AFRSLSALQMTALNKLHHPDYAFGNLSLVVLHNNRTHSLGKCFDGLHSLTLD 240
QY 241 LNNYNDLDFPTAIRTLNLKELGFHNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
Db 241 LNNYNDLDFPTAIRTLNLKELGFHNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
QY 301 FOHLPELRTLNGASQITEFPDLTGTANLESALTGAQISLSPOTVCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLNGASQITEFPDLTGTANLESALTGAQISLSPOTVCNQLPNLQVLDLS 360
QY 361 YNLLLEDPSFVCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIATIHNAFST 420
Db 361 YNLLLEDPSFVCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIATIHNAFST 420
QY 421 LPSLTKLDSNLLSSPFIITGLGTHLKLGNHALQSLISENPELKIEMPIYQCC 480
Db 421 LPSLTKLDSNLLSSPFIITGLGTHLKLGNHALQSLISENPELKIEMPIYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGLWIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGLWIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSSAVLAGVDAFTGSPARHGAWNGVGVCHVIGFLSPASESSVFLTLTAAAL 660
Db 601 AVNMLTGVSSAVLAGVDAFTGSPARHGAWNGVGVCHVIGFLSPASESSVFLTLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMWTIATYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMWTIATYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSNLNLFISPEVTKFILLVVVPLPACILNPLLYILFNPHFKEDIVSLRKQTYV 840
Db 781 PVAFLSPSSNLNLFISPEVTKFILLVVVPLPACILNPLLYILFNPHFKEDIVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVVTFTSSSITVDLPSSVPSPAYPVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVVTFTSSSITVDLPSSVPSPAYPVTTESCHLSS 900
Qy 901 VAFVPCL 907
Db 901 VAFVPCL 907
RESULT 6
ADE59150
ID ADE59150 standard; protein; 907 AA.
XX
AC ADE59150;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein O75473, SEQ ID NO 5041.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
FI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
XX
DR GENBANK; O75473.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDTSLRGVLLSLPVLQLATGGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Qy 61 PSNLSVFTSYLDLSMNNISOLLNPLPSLRFLEELRAGNALTYIPKGAFTGLYSLKVL 120
Db 61 PSNLSVFTSYLDLSMNNISOLLNPLPSLRFLEELRAGNALTYIPKGAFTGLYSLKVL 120
Qy 121 LQNNLRHVPTALQNLRLSLQSLRLDANHSYVPPSCFSGLSHLRLHLDNNALEIPVQ 180
Db 121 LQNNLRHVPTALQNLRLSLQSLRLDANHSYVPPSCFSGLSHLRLHLDNNALEIPVQ 180
Qy 181 AFRSLALQAMTALANKIHHI PDYAFGNLSLVVHLHNNRIHSLGKCKDFGLHSLTLD 240
Db 181 AFRSLALQAMTALANKIHHI PDYAFGNLSLVVHLHNNRIHSLGKCKDFGLHSLTLD 240
Qy 241 LNYNNLDEFFTARTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYNDPIQFVGSA 300
Db 241 LNYNNLDEFFTARTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYNDPIQFVGSA 300
Qy 301 FOHLPELRLTLTLNGASQITTEFPDLTGTLNLESLLTGAQISSLPQTVCNQLQVLDLS 360
Db 301 FOHLPELRLTLTLNGASQITTEFPDLTGTLNLESLLTGAQISSLPQTVCNQLQVLDLS 360
Qy 361 YNLLEDLPSPFVCOQKIDLRHNEIYEIKVDYFQQLLSRLSLNLANWKIAIHPNAPST 420
Db 361 YNLLEDLPSPFVCOQKIDLRHNEIYEIKVDYFQQLLSRLSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLSENPELKVTEMPYAOCC 480
Db 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLSENPELKVTEMPYAOCC 480
Qy 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFQAODERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFQAODERDLEDLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPFKPECHLLDGLWIRIGWTTIAVALTCNALVTSTVFRSPYIYSPIKLLIGVIA 600
Db 541 CSPSPGPFKPECHLLDGLWIRIGWTTIAVALTCNALVTSTVFRSPYIYSPIKLLIGVIA 600
Qy 601 AVNNMLTCVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESVFLLTAAAL 660
Db 601 AVNNMLTCVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESVFLLTAAAL 660
Qy 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMWTIATYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMWTIATYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780

QY 781 PVAFSLFSSLINLTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQYV 840
 Db 781 PVAFSLFSSLINLTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQYV 840
 QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
 Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
 QY 901 VAFVPECL 907
 Db 901 VAFVPECL 907

RESULT 7

AD59153
 ID AD59153 standard; protein; 907 AA.

XX AC AD59153;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein O75473, SEQ ID NO 5044.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GENO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX PR WPI; 2003-268312/26.

XX DR GENBANK; O75473.

XX PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the patented specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 7; Length 907;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDTSLRGVLLSLPVLQLATGSSPRSCVLLRGCPHCHCEPDGRLMLRVDCSDLGLSEL 60
 Db 1 MDTSLRGVLLSLPVLQLATGSSPRSCVLLRGCPHCHCEPDGRLMLRVDCSDLGLSEL 60
 QY 61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
 Db 61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
 QY 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSCFSGLSLRLHLDNNAITFIPVQ 180
 Db 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSCFSGLSLRLHLDNNAITFIPVQ 180
 QY 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVLLHNNRHISLGGKCFGLHSLFTLD 240
 Db 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVLLHNNRHISLGGKCFGLHSLFTLD 240
 QY 241 LNYNNLDEFFTAIRTLNKLGFHNSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
 Db 241 LNYNNLDEFFTAIRTLNKLGFHNSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
 QY 301 FOHLPELRTLNGASQITTEPDLTGTAANLSSLTGTGAQISLSLPOTVCNQLPNQLVDLS 360
 Db 301 FOHLPELRTLNGASQITTEPDLTGTAANLSSLTGTGAQISLSLPOTVCNQLPNQLVDLS 360
 QY 361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAIHHNAPST 420
 Db 361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAIHHNAPST 420
 QY 421 LPSLIKLDLSNLLSSPPIITGLHGLTHLKTGNHALQSLISSENPPELKIEMPAVOCC 480
 Db 421 LPSLIKLDLSNLLSSPPIITGLHGLTHLKTGNHALQSLISSENPPELKIEMPAVOCC 480
 QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEFLDDEEDLKALHSVQ 540
 Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEFLDDEEDLKALHSVQ 540
 QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Db 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 QY 601 AVNMLTGVSNAVLGADAFTEGSPARHGWGCHVIGFLSIFASESSVFLTLTAAAL 660
 Db 601 AVNMLTGVSNAVLGADAFTEGSPARHGWGCHVIGFLSIFASESSVFLTLTAAAL 660
 QY 661 ERGFSVKYSAKFETKAPFSSLSKVIILICALLALTAAPVLLGGSKYGASPLCLPFGGP 720
 Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILICALLALTAAPVLLGGSKYGASPLCLPFGGP 720
 QY 721 STMGYMWALILLNSLCFLMWTIAYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
 Db 721 STMGYMWALILLNSLCFLMWTIAYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
 QY 781 PVAFSLFSSLINLTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQYV 840
 Db 781 PVAFSLFSSLINLTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQYV 840
 QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
 Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900

Qy	901	VAFPCL 907	Db	181	AFRSLALQAWTLALNKIHHPDVFAGNLSLVLVHLHNNRHSGLKKCPDGLHSLETLD	240
Db	901	VAFPCL 907	Qy	241	LNYNLDEPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIQFVGRSA	300
			Db	241	LNYNLDEPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIQFVGRSA	300
			Qy	301	FOHLPRLTLTLNGASQITTEFPDITGTANLESFLTGAQISSLPQTVCNQLPNLQVLDLS	360
			Db	301	FOHLPRLTLTLNGASQITTEFPDITGTANLESFLTGAQISSLPQTVCNQLPNLQVLDLS	360
			Qy	361	YNLLEDLPFSFVCOKLQKIDLRHNEIYEIKVDVTFQQLLSRLSLNLANWKIAIHPNAPST	420
			Db	361	YNLLEDLPFSFVCOKLQKIDLRHNEIYEIKVDVTFQQLLSRLSLNLANWKIAIHPNAPST	420
			Qy	421	LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLSSSENFPELKVTEMPIAYQCC	480
			Db	421	LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLSSSENFPELKVTEMPIAYQCC	480
			Qy	481	AFGCENAYKISNOWNKGDNSMDLHKDAGMFQADDERDLEDFLDFEEDLKALHSVQ	540
			Db	481	AFGCENAYKISNOWNKGDNSMDLHKDAGMFQADDERDLEDFLDFEEDLKALHSVQ	540
			Qy	541	CSPSPGPFKCEHLLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
			Db	541	CSPSPGPFKCEHLLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
			Qy	601	AVNMLTGVSASVLAGVDAFTFGSPARHGANWENGCHVIGFLSIFASESVFLLTLAAL	660
			Db	601	AVNMLTGVSASVLAGVDAFTFGSPARHGANWENGCHVIGFLSIFASESVFLLTLAAL	660
			Qy	661	ERGSVKYSKAFETKAPFSSIKVILLCALLATMAAVPLLGSKYKASPLCLPLPGE	720
			Db	661	ERGSVKYSKAFETKAPFSSIKVILLCALLATMAAVPLLGSKYKASPLCLPLPGE	720
			Qy	721	STMGYVALILLNSLCLFLMNTIATYKLYCNLDKGDLENIDWDCSMVKHIALLLFTNCILNC	780
			Db	721	STMGYVALILLNSLCLFLMNTIATYKLYCNLDKGDLENIDWDCSMVKHIALLLFTNCILNC	780
			Qy	781	PVAFLSFSSILNLTFFISPEVVKFILLVVPPLACINPLLYLTFNPHKEDLVSLRKQTYV	840
			Db	781	PVAFLSFSSILNLTFFISPEVVKFILLVVPPLACINPLLYLTFNPHKEDLVSLRKQTYV	840
			Qy	841	WTRSKHPSLMSINSDDVEKOSCDSTQALVTFTSSSIYDLPSSVPSPAYPVETESCHLSS	900
			Db	841	WTRSKHPSLMSINSDDVEKOSCDSTQALVTFTSSSIYDLPSSVPSPAYPVETESCHLSS	900
			Qy	901	VAFPCL 907	
			Db	901	VAFPCL 907	
					RESULT 9	
					ADH14256	
					ID ADH14256 standard; protein; 907 AA.	
					XX	
					AC ADH14256;	
					XX	
					DT 11-MAR-2004 (first entry)	
					XX	
					DE Human HG38.	
					XX	
					KW human; non-endogenous; G protein-coupled receptor; GPCR; receptor.	
					XX	
					OS Homo sapiens.	
					XX	
					FN US2003105292-A1.	
					XX	
					PD 05-JUN-2003.	
					XX	
					PF 20-SEP-2002; 2002US-00251385.	
					XX	
					PR 26-JUN-1998; 98US-0090783P.	
					PR 07-AUG-1998; 98US-0095677P.	

Qy 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 8
ADG42628
ID ADG42628 standard; protein; 907 AA.
XX
AC ADG42628;
DT 26-FEB-2004 (first entry)
XX
DE Human G protein-coupled receptor 49.
XX
KW cytostatic; gene therapy; NOVX-agonist; NOVX-antagonist; pharmaceutical;
KW NOVX-associated disorder; cancer; human; G protein coupled receptor 49.
XX
OS Homo sapiens.
FN US2003204052-A1.
XX
PD 30-OCT-2003.
XX
PF 04-OCT-2001; 2001US-00970944.
XX
PR 04-OCT-2000; 2000US-0237862P.
XX
PA (HERR/) HERRMANN J L.
PA (RAST/) RASTELLI L.
PA (SHIM/) SHIMKETS R A.
XX
PI Herrmann JL, Rastelli L, Shimkets RA;
XX
DR WPI; 2003-900673/82.
XX
PT New NOVX gene or NOVX-specific antibody, useful for preparing a
PT composition for treating or preventing a NOVX-associated disorder, e.g.,
PT cancer.
XX
PS Disclosure; SEQ ID NO 26; 118pp; English.
XX
CC The invention describes a new isolated polypeptide comprising: a
CC polypeptide or its mature form comprising a sequence not given in the
CC specification; or a variant of (A), where one or more amino acid residues
CC in the variant differs in no more than 15% from the amino acid sequence
CC of the mature form. The pharmaceutical composition may be administered
CC via oral, transdermal, rectal or parenteral route. The polypeptide,
CC nucleic acid or antibody is useful for preparing a composition for
CC treating or preventing a NOVX-associated disorder, e.g., cancer. This is
CC the amino acid sequence of a transmembrane receptor homologue used in a
CC comparison with the novel human proteins of the invention.
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGSGSPRGVLLRGCTHCEPDGRLMLRVDCSDGLSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGSGSPRGVLLRGCTHCEPDGRLMLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYLDLSNMNITQLLPNPLPSLRFLELRAGNALTYPKGAFTGLYSKVLIM 120
Db 61 PSNLSVFTSYLDLSNMNITQLLPNPLPSLRFLELRAGNALTYPKGAFTGLYSKVLIM 120

Qy 121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVPPSCFSGLSLRLHMLDNLALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVPPSCFSGLSLRLHMLDNLALTEIPVQ 180

Qy 181 AFRSLALQAWTLALNKIHHPDVFAGNLSLVLVHLHNNRHSGLKKCPDGLHSLETLD 240


```

PR 13-OCT-1998; 98US-00170496.
XX (LIAW/) LIAW C W.
PA (BEHA/) BEHAN D P.
PA (CHAL/) CHALMERS D T.
XX
PI Liaw CW, Behan DP, Chalmers DT;
XX
DR WPI; 2003-801247/75.
DR N-PSDB; ADH14255.
XX
PT New constitutively active, non-endogenous version of an endogenous human
PT G protein-coupled receptor for the identification of therapeutic
PT compounds, such as agonists.
XX
PS Example 1; SEQ ID NO 264; 227bp; English.
XX
CC The invention relates to a constitutively active, non-endogenous version
CC of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC used for screening therapeutic compounds as inverse agonists, agonists or
CC partial agonists. The GPCR can be also be used to elucidate and
CC understand the roles of GPCRs in normal and diseased humans. The GPCR
CC need not be purified and isolated to be used to screen for therapeutic
CC compounds. The utility of the GPCR as a research tool is enhanced because
CC the role of a particular receptor can be understood before the endogenous
CC ligand is identified. The present sequence is used in the exemplification
CC of the present invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLGVLISLPVLLQATGSSPRSGVLARGCPHCHCEPDGRLRVDCSDGLSEL 60
DB 1 MDTSLGVLISLPVLLQATGSSPRSGVLARGCPHCHCEPDGRLRVDCSDGLSEL 60
QY 61 PNLNVFTSYLDLNMNNTSQQLLPNLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV 120
DB 61 PNLNVFTSYLDLNMNNTSQQLLPNLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVVPSCFSGLSLRLHMLDDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVVPSCFSGLSLRLHMLDDNALTEIPVQ 180
QY 181 APRSLAQAMTALNKHIIHIDYAFGNLSLVLHNNRIHSLGKCFDGLHSLDTLD 240
DB 181 APRSLAQAMTALNKHIIHIDYAFGNLSLVLHNNRIHSLGKCFDGLHSLDTLD 240
QY 241 LNNYNDLDFPTAIRLSNLKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPVQVGRSA 300
DB 241 LNNYNDLDFPTAIRLSNLKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPVQVGRSA 300
QY 301 FOHLPELRLTLNGASQITEFPDLTGATNLESITLTGAQISSLPQVCNQLPNLQVLDLS 360
DB 301 FOHLPELRLTLNGASQITEFPDLTGATNLESITLTGAQISSLPQVCNQLPNLQVLDLS 360
QY 361 YNLLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIALIHNAFST 420
DB 361 YNLLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIALIHNAFST 420
QY 421 LPSLKLDSLSSNLSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
DB 421 LPSLKLDSLSSNLSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNQDNSSMDLHKDAGMFOAQDERDLEDLFDPEEDLKALHSVQ 540
DB 481 AFGVCENAYKISQNWKNQDNSSMDLHKDAGMFOAQDERDLEDLFDPEEDLKALHSVQ 540
QY 541 CSPSPGPFKPCSHLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
DB 541 CSPSPGPFKPCSHLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDATFTGSPARHGAWWENGCHVIGFLSIFASESSVFLTLTAA 660
DB 601 AVNMLTGVSASVLAGVDATFTGSPARHGAWWENGCHVIGFLSIFASESSVFLTLTAA 660
QY 661 ERGFSVKYSKAFETKAPESLSKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGEP 720
DB 661 ERGFSVKYSKAFETKAPESLSKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGEP 720
QY 721 STMGYVVALIILNSLCFLMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYVVALIILNSLCFLMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTV 840
DB 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITYDLPSPSSVPSPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITYDLPSPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

```

RESULT 10

ADN40013

ID ADN40013 standard; protein; 907 AA.

XX ADN40013;

AC ADN40013;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C383.

Human; differential expression; cancer; angiogenic disorder;
 fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 inflammatory disease; autoimmune disease;
 retinal neovascularisation syndrome; scarring; uterine fibroid;
 detection; diagnosis; prognosis; drug screening; drug targeting;
 wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 vulnerability; gene therapy; vaccine.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

XX 21-NOV-2001; 2001US-0332464P.

XX 29-NOV-2001; 2001US-0334393P.

XX 03-DEC-2001; 2001US-0335394P.

XX 14-DEC-2001; 2001US-0340376P.

XX 08-JAN-2002; 2002US-0347211P.

XX 10-JAN-2002; 2002US-0347349P.

XX 08-FEB-2002; 2002US-0355250P.

XX 13-FEB-2002; 2002US-0356714P.

XX 20-FEB-2002; 2002US-0359077P.

XX 29-MAR-2002; 2002US-0368809P.

XX 04-APR-2002; 2002US-0370110P.

XX 12-APR-2002; 2002US-0372246P.

XX 05-JUN-2002; 2002US-0386614P.

XX 16-JUL-2002; 2002US-0396839P.

XX 22-JUL-2002; 2002US-039775P.

XX 22-JUL-2002; 2002US-0397845P.

XX 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PA

XX

PI Afaf D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DR, Murray R, Watson SR, Wilson KE, Ziolknik A;
DR WPI; 2003-468649/44.
DR N-PSDB; ADN39796.

PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C383; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.

XX Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60

Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLFLLEELRAGNALTYPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLFLLEELRAGNALTYPKGAFTGLYSKVLV 120

Qy 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180

Qy 181 AFRSLALQAMTLALNKTHHPDYAFGNLSLVLHLHNNRIHSLGKKCFDGLHSLTLD 240
Db 181 AFRSLALQAMTLALNKTHHPDYAFGNLSLVLHLHNNRIHSLGKKCFDGLHSLTLD 240

Qy 241 LNYNNLDEFPPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLITIHFDYDNP10FVGRSA 300
Db 241 LNYNNLDEFPPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLITIHFDYDNP10FVGRSA 300

Qy 301 FQHLPELRTLNGASQITTEPDLTGTANLESILTGAQISSLPQTVCNQLPNQLVLDLS 360
Db 301 FQHLPELRTLNGASQITTEPDLTGTANLESILTGAQISSLPQTVCNQLPNQLVLDLS 360

Qy 361 YNLLEDLPFSVCQKQKIDLRHNEIYKVDYTFQQLLSLRLSLNLANWKIAIHPNAFST 420
Db 361 YNLLEDLPFSVCQKQKIDLRHNEIYKVDYTFQQLLSLRLSLNLANWKIAIHPNAFST 420

Qy 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPPELKIEMPYAYOC 480
Db 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPPELKIEMPYAYOC 480

Qy 481 AFGVCENAYKISNQWKNKDNSSMDLHKDKAGMQAQRDLFDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNQWKNKDNSSMDLHKDKAGMQAQRDLFDLDFEEDLKALHSVQ 540

Qy 541 CSPSPGPFKCEHLLDGLIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db 541 CSPSPGPFKCEHLLDGLIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Qy 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWENGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWENGCHVIGFLSIFASESSVFLTLAAL 660

Qy 661 ERGFSVKYSKAFETKAPFSSIKVILLACALLATMAAVPLLGSKYKASPLCLPLPGE 720
Db 661 ERGFSVKYSKAFETKAPFSSIKVILLACALLATMAAVPLLGSKYKASPLCLPLPGE 720

Qy 721 STMGYVALIILNLSLCLFLMTIATKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNLSLCLFLMTIATKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780

Qy 781 PVAFLSFSSILNLTIFISPEVIFKILLVVVPLPACINPLLYLILFNPHFKEDLVSRKQTY 840
Db 781 PVAFLSFSSILNLTIFISPEVIFKILLVVVPLPACINPLLYLILFNPHFKEDLVSRKQTY 840

Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 11
ADN39531
ID ADN39531 standard; protein; 907 AA.
XX AC ADN39531;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A131.
XX KW Human; differential expression; cancer; angiogenic disorder;
XX KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX KW inflammatory disease; autoimmune disease;
XX KW retinal neovascularisation syndrome; scarring; uterine fibroid;
XX KW detection; diagnosis; prognosis; drug screening; drug targeting;
XX KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX KW vulnary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
XX 29-NOV-2001; 2001US-0332464P.
XX 03-DEC-2001; 2001US-0334393P.
XX 14-DEC-2001; 2001US-0335394P.
XX 08-JAN-2002; 2001US-0340376P.
XX 10-JAN-2002; 2002US-0347211P.
XX 08-FEB-2002; 2002US-0347349P.
XX 13-FEB-2002; 2002US-035250P.
XX 20-FEB-2002; 2002US-0356714P.
XX 29-MAR-2002; 2002US-036809P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-0397775P.
XX 09-SEP-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX

XX	WPI; 2004-390329/36.	361	YNLEDLPSFVSCVKIQKIDLRHNEIYEIKVDYTFQQLLSLSLAWNKIALIHHNAPST	420
DR	N-PSDB; AD029923.	361	YNLEDLPSFVSCVKIQKIDLRHNEIYEIKVDYTFQQLLSLSLAWNKIALIHHNAPST	420
XX				
PT	Novel mammalian G protein coupled receptors, useful for identifying	421	LPSLTKLDSNLSLSSFFITGLHGLTHLKLGNHALQSLISSENFPELKIEMPIYQCC	480
PT	compounds that modulates diagnosing and treating disease condition	421	LPSLTKLDSNLSLSSFFITGLHGLTHLKLGNHALQSLISSENFPELKIEMPIYQCC	480
PT	associated with GPCR dysfunction e.g. autoimmune diseases, angina			
PT	pectoris, Parkinson's disease.			
XX				
PS	Claim 151; SEQ ID NO 510; 542pp; English.	481	AFVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLLDFEEDLKALHSVQ	540
XX		481	AFVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLLDFEEDLKALHSVQ	540
CC	The invention relates to human and mouse G protein-coupled receptors	541	CSPSPGPKCEHLLDGLIRIGVWTIAVLTALTCNALVTSTVFRSPLVISPIKLLIGVIA	600
CC	(GPCRs) and nucleic acids encoding them. The invention also relates to	541	CSPSPGPKCEHLLDGLIRIGVWTIAVLTALTCNALVTSTVFRSPLVISPIKLLIGVIA	600
CC	sequences at least 90% identical to the GPCR proteins and nucleic acids			
CC	of the invention; methods of treating, preventing or diagnosing diseases			
CC	associated with GPCRs of the invention; methods of screening for			
CC	compounds useful in the treatment of GPCR-related diseases; a transgenic			
CC	mouse comprising a GPCR gene of the invention; a mouse comprising a			
CC	mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived			
CC	from the transgenic mice; kits comprising several mice, each of which has			
CC	a mutation in a different GPCR gene of the invention; and kits comprising			
CC	probes which hybridize to GPCR polynucleotides of the invention. The			
CC	invention further discloses variants of the GPCR polypeptides and vectors			
CC	comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may			
CC	be used in the diagnosis, treatment or prevention of a wide variety of			
CC	diseases including neurological disorders (e.g., Alzheimer's disease,			
CC	depression, diabetic neuropathy, Parkinson's disease or schizophrenia);			
CC	disorders of the adrenal gland; disorders of the colon or intestine			
CC	(e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel			
CC	syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or			
CC	myocardial infarction); muscular disorders; blood disorders (e.g.,			
CC	anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or			
CC	AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid			
CC	arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,			
CC	obesity, enzyme deficiency-related diseases or vitamin deficiency-related			
CC	diseases); and disorders of the kidney, liver, lung, breast, ovary,			
CC	uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and			
CC	thyroid (e.g., cancers). The present sequence represents a GPCR of the			
CC	invention. Note: The full sequence data for this patent did not form part			
CC	of the printed specification; those sequences not shown were obtained in			
CC	electronic format directly from WIPO at			
CC	ftp.wipo.int/pub/published_pct_sequences.			
XX				
SQ	Sequence 907 AA;			
	Query Match 100.0%; Score 4702; DB 8; Length 907;			
	Best Local Similarity 100.0%; Pred. No. 0;			
	Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60		
DB	1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60		
QY	61 PNLSVFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYIPKGAFTGLYSLKVL	120		
DB	61 PNLSVFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYIPKGAFTGLYSLKVL	120		
QY	121 LQNNLRHVPTALQNLRSLSQSLRDANHSVVPSPCSGLSLRHLMDNALTEIPVQ	180		
DB	121 LQNNLRHVPTALQNLRSLSQSLRDANHSVVPSPCSGLSLRHLMDNALTEIPVQ	180		
QY	181 AFRSLALQAMTLALNKIHHIDYAFGNLSLVLHLHNNRHSLSGKFCGLHSLETID	240		
DB	181 AFRSLALQAMTLALNKIHHIDYAFGNLSLVLHLHNNRHSLSGKFCGLHSLETID	240		
QY	241 LNYNNLDEPTAIRLSNLSKELGFSNNIRSIPEKAFVGNPSLITIFHYDNPQVGRSA	300		
DB	241 LNYNNLDEPTAIRLSNLSKELGFSNNIRSIPEKAFVGNPSLITIFHYDNPQVGRSA	300		
QY	301 FOHLPELRLTLNGASQITEFPDLTGTLNLSLITGAQISSLPQVNCNQLNQLVDLS	360		
DB	301 FOHLPELRLTLNGASQITEFPDLTGTLNLSLITGAQISSLPQVNCNQLNQLVDLS	360		

QY	361	YNLEDLPSFVSCVKIQKIDLRHNEIYEIKVDYTFQQLLSLSLAWNKIALIHHNAPST	420
DB	361	YNLEDLPSFVSCVKIQKIDLRHNEIYEIKVDYTFQQLLSLSLAWNKIALIHHNAPST	420
QY	421	LPSLTKLDSNLSLSSFFITGLHGLTHLKLGNHALQSLISSENFPELKIEMPIYQCC	480
DB	421	LPSLTKLDSNLSLSSFFITGLHGLTHLKLGNHALQSLISSENFPELKIEMPIYQCC	480
QY	481	AFVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLLDFEEDLKALHSVQ	540
DB	481	AFVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLLDFEEDLKALHSVQ	540
QY	541	CSPSPGPKCEHLLDGLIRIGVWTIAVLTALTCNALVTSTVFRSPLVISPIKLLIGVIA	600
DB	541	CSPSPGPKCEHLLDGLIRIGVWTIAVLTALTCNALVTSTVFRSPLVISPIKLLIGVIA	600
QY	601	AVNMLTGVSAAVLAVDAFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL	660
DB	601	AVNMLTGVSAAVLAVDAFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL	660
QY	661	ERGFVSKYSAKFETKAPSSLKVIILLCALLALTAALTAALTAALTAALTAALTAAL	720
DB	661	ERGFVSKYSAKFETKAPSSLKVIILLCALLALTAALTAALTAALTAALTAALTAAL	720
QY	721	STMGVWVALIILNSLCFLMMTIATKLYCNLDKGLNIDWDCSMVKHIALLLFTNCILNC	780
DB	721	STMGVWVALIILNSLCFLMMTIATKLYCNLDKGLNIDWDCSMVKHIALLLFTNCILNC	780
QY	781	PVAFSLFSSSLNLFISPEVIFKILLVVVPLPACLNPLLYILFNPHFKEDVSLRKQTVV	840
DB	781	PVAFSLFSSSLNLFISPEVIFKILLVVVPLPACLNPLLYILFNPHFKEDVSLRKQTVV	840
QY	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTTESCHLS	900
DB	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTTESCHLS	900
QY	901	VAFPVCL 907	
DB	901	VAFPVCL 907	
RESULT 14			
ADQ80369			
ID	ADQ80369	standard; protein; 907 AA.	
XX			
AC	ADQ80369;		
XX			
DT	21-OCT-2004	(first entry)	
XX			
DE		G protein-coupled receptor 49 protein.	
XX			
KW		Cytostatic; epidermal-growth-factor-receptor modulator; identification;	
KW		therapeutic response; cancer; EGFR; biomarker.	
XX			
OS		Homo sapiens.	
XX			
PN		WO2004063709-A2.	
XX			
PD		29-JUL-2004.	
XX			
PF		08-JAN-2004; 2004WO-US0000368.	
XX			
PR		08-JAN-2003; 2003US-0438735P.	
XX			
PA		(BRIN) BRISTOL-MYERS SQUIBB CO.	
XX			
PI		Amler LC, Januario-T.	
XX			
DR		WPI; 2004-544114/52.	
XX			
DR		N-PSDB; ADQ80249.	
XX			
PT		Identifying a mammal that will respond therapeutically to a method of	
PT		treating cancer comprises the level of a biomarker in a mammal	

PT before and after exposure to an epidermal growth factor receptor (EGFR)
XX modulator.
PS Disclosure; SEQ ID NO 141; 520pp; English.
XX
CC The invention relates to a method of identifying a mammal that will
CC respond therapeutically to a method of treating cancer by administering
CC an epidermal growth factor receptor (EGFR) modulator by comparing the
CC level of a biomarker in a mammal before and after exposure to an EGFR
CC modulator. The method comprises: (a) measuring, in the mammal, the level
CC of at least one biomarker identified in the specification; (b) exposing
CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
CC level of the biomarker, where a difference in the level in step (c)
CC compared to step (a) indicates that the mammal will respond
CC therapeutically to the method of treating cancer. The method and
CC biomarkers are useful for identifying a mammal that will respond
CC therapeutically to a method of treating cancer by administering an
CC epidermal growth factor receptor (EGFR) modulator. This sequence
CC corresponds to one of the biomarkers whose levels of expression is
CC measured in the method of the invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 8; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
DB 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYDLSNNISQLLPPLSLRLEELRAGNALTYIPKGAFTGLSKVLTM 120
DB 61 PSNLSVFTSYDLSNNISQLLPPLSLRLEELRAGNALTYIPKGAFTGLSKVLTM 120
QY 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHSYVPPSCFSGLSRLHLWLDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHSYVPPSCFSGLSRLHLWLDNALTEIPVQ 180
QY 181 AFRSLALQMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKCPDGLHSLETLD 240
DB 181 AFRSLALQMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKCPDGLHSLETLD 240
QY 241 LNYNLDLPEFTAIRTLNKLGHSHNNIRISPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
DB 241 LNYNLDLPEFTAIRTLNKLGHSHNNIRISPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
QY 301 FQHLPELRTLINGASQITTEPDLTGTLNLSLTLTGAQISLPTQVCNQLPQLVLDLS 360
DB 301 FQHLPELRTLINGASQITTEPDLTGTLNLSLTLTGAQISLPTQVCNQLPQLVLDLS 360
QY 361 YNLLEDLPSPVQCKLQKIDLRHNEIYKVDTFQQLLSLRLSLANNNKIAIHPNAPST 420
DB 361 YNLLEDLPSPVQCKLQKIDLRHNEIYKVDTFQQLLSLRLSLANNNKIAIHPNAPST 420
QY 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKGHALQSLISENPELKVLEMPYVQCC 480
DB 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKGHALQSLISENPELKVLEMPYVQCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFEDLKALHSVQ 540
QY 541 CSPSPGPKPEHLLDGLMIRIGVWTIAVLATCNALVTSTVFRSPLYSIPKLLIGVIA 600
DB 541 CSPSPGPKPEHLLDGLMIRIGVWTIAVLATCNALVTSTVFRSPLYSIPKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSGFARHGAWENGVCVIGFLSTFASSESVFLTLAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSGFARHGAWENGVCVIGFLSTFASSESVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPFSSLKVILLCALLALTAAPVLLGGSKYGASPLCLPLPFGEP 720

DB 661 ERGFSVKYSAKFETKAPFSSLKVILLCALLALTAAPVLLGGSKYGASPLCLPLPFGEP 720
QY 721 STMGYVALILLNSLCLFMMMTIATYKLYCNLDKGDLENIWDCSMVKHIALLEFNCILNC 780
DB 721 STMGYVALILLNSLCLFMMMTIATYKLYCNLDKGDLENIWDCSMVKHIALLEFNCILNC 780
QY 781 PVAFLSPSSLINLTFISPEVIKFIILLVVPVLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
DB 781 PVAFLSPSSLINLTFISPEVIKFIILLVVPVLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFSTSSITVDLPPSSVPSPAYPVVTSCHLSS 900
DB 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFSTSSITVDLPPSSVPSPAYPVVTSCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907
RESULT 15
ADR67868
ID ADR67868 standard; protein; 907 AA.
XX
XX ADR67868;
AC
XX 18-NOV-2004 (first entry)
DT
XX Human HG38 protein.
DE
XX human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
KW
XX Homo sapiens.
OS
XX WO2004074436-A2.
FN
XX 02-SEP-2004.
PD
XX 11-FEB-2004; 2004WO-US004060.
PF
XX 19-FEB-2003; 2003US-0448959P.
PR
XX (INCY-) INCYTE CORP.
PA
XX Lasek AW;
PI
XX WPI; 2004-652946/63.
DR
XX N-PSDB; ADR67869.
XX
XX Detecting colon or lung cancer, by determining amount of protein in
XX sample, comparing amount of protein to standard, and differential
XX expression of protein in sample indicates colon or lung cancer.
XX
XX Claim 4; SEQ ID NO 1; 79pp; English.
XX
XX This sequence represents the human G protein-coupled receptor (GPCR)
XX known as HG38. The HG38 protein and corresponding nucleic acid, may be
XX used in the method of the invention for detecting colon or lung cancer.
XX The first method involves performing an assay to determine the amount of
XX HG38 in a sample of colon or lung tissue, and comparing the amount of
XX protein to standard, thus detecting expression of protein in sample,
XX where differential expression of protein in sample when compared with the
XX standard is diagnostic of colon or lung cancer. The second method
XX involves hybridizing a composition comprising the HG38 coding sequence,
XX or its complement, and a labelling moiety, to nucleic acids of a sample,
XX of colon or lung tissue under conditions to form at least one
XX hybridization complex, detecting hybridization complex formation, and
XX comparing complex formation of the polynucleotide in the sample relative to
XX differential expression of the polynucleotide in the sample relative to
XX the standard and is diagnostic of a colon or lung cancer. This first
XX method enables earlier diagnosis before the patient is symptomatic. Anti-
XX HG38 antibodies are useful for treating colon or lung cancer.
XX
XX Sequence 907 AA;

	Query Match	100.0%;	Score 4702;	DB 8;	Length 907;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 907;	Conservative	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MDTSRLGVLLSLPVLQLATCGSPRSQVLLRGCPHCHCEPDGRMLRVDCSDLGLSEL	60		
Db	1	MDTSRLGVLLSLPVLQLATCGSPRSQVLLRGCPHCHCEPDGRMLRVDCSDLGLSEL	60		
Qy	61	PSNLSVFTSYLDLSNNNISQLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSLKVL	120		
Db	61	PSNLSVFTSYLDLSNNNISQLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSLKVL	120		
Qy	121	LQNNQLRHVPTEALQNLRSLOSLRDANHI SVVPPSCFSGLHSLRHLMDNALTEIPVQ	180		
Db	121	LQNNQLRHVPTEALQNLRSLOSLRDANHI SVVPPSCFSGLHSLRHLMDNALTEIPVQ	180		
Qy	181	AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRIHSLGKKCFDGLHSLETL	240		
Db	181	AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRIHSLGKKCFDGLHSLETL	240		
Qy	241	LNYNLDEFPPTAIRTLNKLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQIPVGRSA	300		
Db	241	LNYNLDEFPPTAIRTLNKLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQIPVGRSA	300		
Qy	301	FOHLPELRTLNLGASQITTEPDDLTGTANLESLITGTGAQISSLPOTVCNQLPNLQVLDLS	360		
Db	301	FOHLPELRTLNLGASQITTEPDDLTGTANLESLITGTGAQISSLPOTVCNQLPNLQVLDLS	360		
Qy	361	YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAI IHENAFST	420		
Db	361	YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAI IHENAFST	420		
Qy	421	LPSLIKLDLSNLLSSPITGLHGLTHLKTGNHALQSLIISSENPPELKVTEMPYQCC	480		
Db	421	LPSLIKLDLSNLLSSPITGLHGLTHLKTGNHALQSLIISSENPPELKVTEMPYQCC	480		
Qy	481	AFVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ	540		
Db	481	AFVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ	540		
Qy	541	CSPSPGPKPCHELLDGLNLRIGVWTVIAVLATCNALVTSTVFRSPLYISPIKLIGVIA	600		
Db	541	CSPSPGPKPCHELLDGLNLRIGVWTVIAVLATCNALVTSTVFRSPLYISPIKLIGVIA	600		
Qy	601	AVNMLTGVSASVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAL	660		
Db	601	AVNMLTGVSASVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAL	660		
Qy	661	ERGFSVKYSAKFETKAPFSSLKVIILLCALLATAAAVPLLGSKYKASPLCLPLPFGEP	720		
Db	661	ERGFSVKYSAKFETKAPFSSLKVIILLCALLATAAAVPLLGSKYKASPLCLPLPFGEP	720		
Qy	721	STMGYVVALILLNSLCLFLMMTIAVTKYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Db	721	STMGYVVALILLNSLCLFLMMTIAVTKYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Qy	781	PVAFPLSPSSLINLTFI SPEVTKFILLVVVPLPACLNPLLYTLFNPHFKEDLVSLRKQTV	840		
Db	781	PVAFPLSPSSLINLTFI SPEVTKFILLVVVPLPACLNPLLYTLFNPHFKEDLVSLRKQTV	840		
Qy	841	WTRSKHPSLMSINSDVVEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVFTESCHLSS	900		
Db	841	WTRSKHPSLMSINSDVVEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVFTESCHLSS	900		
Qy	901	VAFPVCL 907			
Db	901	VAFPVCL 907			

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:30 ; Search time 104 Seconds
(without alignments)
4465.921 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGLVLLSLPVLQLAT.....PAYPVTSCHLSLVAFPCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4702	100.0	907	1 LGR5 HUMAN	O75473 homo sapien
2	4048	86.1	907	1 LGR5 MOUSE	Q9z1p4 mus musculus
3	2492.5	53.0	928	2 Q9BYD7	Q9byd7 homo sapien
4	2412.5	51.3	923	2 Q86VU0	Q86vu0 homo sapien
5	2410	51.3	915	2 Q6UY15	Q6uy15 homo sapien
6	2091.5	44.5	951	1 LGR4 HUMAN	Q9bxb1 homo sapien
7	2088.5	44.4	951	1 LGR4 RAT	Q9z2h4 rattus norv
8	2024.5	43.1	927	2 Q8N537	Q8n537 homo sapien
9	1980	42.1	828	1 LGR6 HUMAN	Q9hbx8 homo sapien
10	1942.5	41.3	878	2 Q8BX59	Q8bx59 mus musculus
11	1508	34.2	363	2 Q8C8A7	Q8c8a7 mus musculus
12	1182.5	25.1	1093	2 Q8HA06	Q8ha06 crassostrea
13	1109.5	23.6	1257	2 Q7PNF8	Q7pnf8 anopheles g
14	1070	22.8	1360	2 Q7KTA0	Q7kta0 drosophila
15	1068	22.7	1360	2 Q9NDI1	Q9ndi1 drosophila
16	1058.5	22.5	1012	2 Q95YI6	Q95yi6 asterina pe
17	1058.5	22.5	1280	2 Q95YI7	Q95yi7 asterina pe
18	976	20.8	1300	2 Q9NKD6	Q9nk6 drosophila
19	947.5	20.2	459	2 Q8R301	Q8r301 mus musculus
20	902.5	19.2	1050	2 Q9BN18	Q9bn18 drosophila
21	834.5	17.7	488	2 Q6PHA3	Q6pha3 mus musculus
22	829.5	17.6	1039	2 Q86BL1	Q86bl1 drosophila
23	695.5	14.8	412	2 Q8BZR7	Q8bzr7 mus musculus
24	656.5	14.0	692	1 FSHR MOUSE	P35378 mus musculus
25	655	13.9	695	1 FSHR PIG	P49059 sus scrofa
26	650	13.8	695	1 FSHR BOVIN	P35376 bos taurus
27	648.5	13.8	695	1 FSHR MACFA	P32212 macaca fasc
28	642.5	13.7	694	1 FSHR HORSE	P47799 equus cabal
29	642	13.7	695	1 FSHR SHEEP	P35379 ovis aries
30	636	13.5	692	1 FSHR RAT	P20395 rattus norv
31	634.5	13.5	694	2 Q6YNB6	Q6ynb6 macropus eu

32	633	13.5	687	1 FSHR EQUAS	Q95179 equus asinu
33	625.5	13.3	693	2 Q7ZTV5	Q7ztv5 cairina mos
34	625	13.3	688	2 Q64183	Q64183 rattus sp.
35	625	13.3	696	2 Q9DGF5	Q9dgf5 cynops pyri
36	622	13.2	701	2 Q9DGC6	Q9dgc6 oreochromis
37	622	13.2	710	2 Q8AXM5	Q8axm5 clarias gar
38	621.5	13.2	695	1 FSHR HUMAN	P23945 homo sapien
39	621	13.2	147	2 Q7IM96	Q7im96 sus scrofa
40	616	13.1	694	2 Q6R6L8	Q6r6l8 mesocricetu
41	616	13.1	695	2 Q8R428	Q8r428 cavia porce
42	606.5	12.9	704	2 Q6POA4	Q6poa4 sparus auru
43	606.5	12.9	708	2 Q6TCF8	Q6tcf8 brachydanio
44	600	12.8	700	1 LSHR MOUSE	P30730 mus musculu
45	594	12.6	693	1 FSHR_CHICK	P79763 gallus gall

ALIGNMENTS

RESULT 1

ID LGR5 HUMAN STANDARD; PRT; 907 AA.

AC O75473; Q9UP75;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor 49)

DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor 49)

GN Name=GPR49; Synonyms=GPR67, LGR5;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RX MEDLINE=98308104; PubMed=9642114; DOI=10.1006/bbrc.1998.8774;

RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;

RT "Identification and cloning of an orphan G protein-coupled receptor of the glycoprotein hormone receptor subfamily."

RL Biochem. Biophys. Res. Commun. 247:266-270(1998).

RN [2]

SEQUENCE FROM N.A.

RP TISSUE=Placenta;

RC MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;

RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;

RT "Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region."

RL Mol. Endocrinol. 12:1830-1845(1998).

CC -!- FUNCTION: Orphan receptor. It may be an important receptor for signals controlling growth and differentiation of specific embryonic tissues (by similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal cord, and various region of brain.

CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC -!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF062006; AAC28019.1; -

CC EMBL; AF061444; AAC77911.1; -

DR PIR; JE0176; JE0176.

DR HSSP; Q9BZR6; IOZN.

DR Genew; HGNC:4504; GPR49.

DR MIM; 606667; -


```

Qy 661 ERGSVKYSAKPTKAPFSSKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPGEF 720
Db 661 ERGSVKYSAKPTKAPFSSKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPGEF 720
Qy 721 STMGYVALIILLNSLCFLMMTIAATKGYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILLNSLCFLMMTIAATKGYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSSNLNLTISPSEVIFKFIILVVVPLPACINPLLYILFNPHEKEDVLSLRKQTYV 840
Db 781 PVAFLSFSSNLNLTISPSEVIFKFIILVVVPLPACINPLLYILFNPHEKEDVLSLRKQTYV 840
Qy 841 WTRSKHPSLMSINDVDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINDVDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 3
Q9BYD7
ID Q9BYD7 PRELIMINARY; PRT; 928 AA.
AC Q9BYD7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE VTS20631 (Fragment).
GN Name=VTS20631;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki H., Havaashi A., Kozuma S., Saito T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049405; BAB39854.1; -.
DR HSSP; P25147; 1D08.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmu_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON_TER 1
SQ SEQUENCE 928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;

Query Match 53.0%; Score 2492.5; DB 2; Length 928;
Best Local Similarity 55.7%; Pred. No. 3.4e-140;
Matches 491; Conservative 134; Mismatches 226; Indels 31; Gaps 9;

Qy 39 HCEPDGMLLRVDCSDGLSELPSNLVFTSYDLSMNNISQLLPNPLPSLRFLEELRLA 98
Db 1 HCEPDG-IMLSADCSLGLSAVPGDLPLTAYLDLSMNNLTQELQGLFHHLRLEELRLS 59

Qy 99 GNALTYIPKGAFTGLYSKVLMLQNNQRLRHVPTAEALQNLRLSLQSLRLDANHSIYVPPSCF 158
Db 60 GNHLSHIPQAFSGLYSLKILMQNNQGLGIPAEALWELPSLQSLRLDANHSIYVPPSCF 119

Qy 159 SCLSHSLHLLDDNALTPIVQAFSLAQMTALANKHHPIDYAFGNLSLVLHLH 218
Db 120 EGLSLSHLHLLDDNALTPIVQAFSLAQMTALANKHHPIDYAFGNLSLVLHLH 179

Qy 219 NNRHSLGKCKFDGLHSLSETLDLNNYNNDEPPTAIRTLSNLKELGFHSHNNIRSIPEKAFV 278

```

```

Db 180 NNRHQLGTHSFEGHLNLTLDLNNYNLQBFPPVAIRTGLRQLQELGFHNNNIKAIPKAFM 239
Qy 279 GNPSLIITHFVDNPIQFVGRSAFOHLPELRLTLTLNGASQITERFDLGTANLESITLGA 338
Db 240 GNPLQLTIHFVDNPIQFVGRSAFOYLPKHLHTLSLNGAMDIOEFPDLGKTTSLLEILTTRA 299
Qy 339 QISSLPQTVCNQPNLQVLDLSYNLLEDPSFSVCQKLDKIDLRHNEIYEIKVDTFQOLL 398
Db 300 GIRLLPSGMCOQLPRLVLELSHNOIEELPSLHRCQKLEIGLQHNRIWEIGADTFQSLS 359
Qy 399 SLRSNLAWNKIAIHPNAPSTLPSLKLDLSSNLLSPFTTGLHGLTHLKLTKGNHALQS 458
Db 360 SIQALDLSWNAIRSIHPPEASTLHSLVKLDLTNQLTTLPLAGLGLMHLKLGKLNALSQ 419
Qy 459 LLSSENPELVKVIEMPYAYOCCAFGCENAYKISNONKNGDNSSMDLH-----KKDA 511
Db 420 AFSKDSFPKRLILEVPYAYOCCPYGMCASFPKASGOW-----BAEDLHLDDESSKRPL 473
Qy 512 GMFOAQDB----RDELDLDFEEDLKALHSVQCSPPSPGPKCEHLLDGLWLRIGVWTI 567
Db 474 GLLAQAEHNDQDLDELQLEM-EDSKPHSPVQCSPTFGPKCEYLFEWGIRLAVWAI 532
Qy 568 AVLATCNALVTSTVFR-SPLYISPIKLLIGVIAAVNMLTGVSSAVLAGVDAFTFGSPAR 626
Db 533 VLLSVLCNGLVLLTVFAGGPAPLPVFKVVGAIAGANTLTGISCGLLASVDALTFGQFSE 592
Qy 627 HGAMWENGVCCHVIGLSIFASESSVLLTLAALRGFSVKYSKPFETKAPFSSIKVILL 686
Db 593 YGARWETGLGCRATGFLAVLGSEASVILLTLAAVQCSVSCVRAYGKPSGLSVRAGVL 652
Qy 687 LCALLALTMAAVPLLGSKYKASPLCLPL--PFGEPTMGYMWVALIILLNSLCFLMMTIA 744
Db 653 GCLALAGLAALPLASVEYKASPLCLPYAPPEQPAALGFTVALVMNNSFCFLVAVAG 712
Qy 745 TKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNCPVAFISFSSLINLTISPSEVIFKFI 804
Db 713 IKLYCDLPRGDFEAVWDCAMVRHVAWLIFADGLLYCFVAFISFASMLGLFPVTEAVKSV 772
Qy 805 LLVVPLPACINPLLYILFNPHEKEDVLSLRKQTYVWTRSKHPSLMSINDVDEKQSCDS 864
Db 773 LLVVPLPACINPLLYILFNPHEKEDVLSLRKQTYVWTRSKHPSLMSINDVDEKQSCDS 827
Qy 865 TQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSSVAFVPC 906
Db 828 TQALVAFSDVDLILEASEAGRP----FGLTYVGFPSVTLSLC 865

RESULT 4
Q86VU0
ID Q86VU0 PRELIMINARY; PRT; 923 AA.
AC Q86VU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=IcR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood.
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buétow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.L., Skalska U., Smallus D.E., Schmerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC047905; AAH47905.1; -;
DR HSSP: Q9BZE6; 1P8T
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR Rhodpsn.
DR InterPro: IPR002131; Gphrmn_receptor.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam: PF00560; LRR_1; 15.
DR PRINTS: PR00373; GLYCHORMONER.
DR PRINTS: PR00237; GPCR RHODOPSIN.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00369; LRR_TYP; 8.
FT NON TER 1
SQ SEQUENCE 923 AA; 100047 MW; 541D6746DAB06813 CRC64;
Query Match 51.3%; Score 2412.5; DB 2; Length 923;
Best Local Similarity 54.3%; Pred. No. 2e-135;
Matches 483; Conservative 133; Mismatches 230; Indels 43; Gaps 10;
QY 32 RGCPTHCCEPDGRMLRLVDCSDGLSELPSNLVFTSYLDLSNMNLSQLLPNLPISLRF 91
DB 1 KGKPTG---QGMGRPRLTLCV-----QVSIISARDLSMNLTQLPGLPHLRF 47
QY 92 LEEELAGNALTYPKGAFTGLYSKVLMLQNNQRHVPTEALQNLRSLSRLDANHS 151
DB 48 LEEELSGNHLSHIPQAFSGLYSKILMLQNNQGGIPAEALWELPSQLSLRLDANHS 107
QY 152 YVPPSCFSLSLRLHLDNLTALTEIPVQAPSLSALQAMTLAKNIHHPDYAFGNLSS 211
DB 108 LVPERSFGLSLRLHLDNLTALTEIPVRLNLPALQAMTLAKNIHHPDYAFGNLSS 167
QY 212 LVVLHLHNRNTHSLGKCFDGLHSLTETLDLNTNLDLDEFPPTAIRTLNKLKELGFTSNNTRS 271
DB 168 LVVLHLHNRNTHSLGKCFDGLHSLTETLDLNTNLDLDEFPPTAIRTLNKLKELGFTSNNTRS 227
QY 272 IPEKAFVGNPSLIITHFYDNPITQFVGRSAFQHLPELRLTLNGASQITEFPDPLTGANLE 331
DB 228 IPEKAFVGNPSLIITHFYDNPITQFVGRSAFQHLPELRLTLNGASQITEFPDPLTGANLE 287
QY 332 SILTGAOISSIPQVCHQNLNQLVDLSYNLEDLPSFSVCOKLQKLDLRHNEIYELKV 391
DB 288 ILTLTRAGIRLLPSGMCQQLPRLRLVLESHNQIEELPSLHRCQKLEELGLQNRITWEIGA 347
QY 392 DTFOQLLSRLSLNLAWNKIAIHPNFAFSLPSLIKLDLSNLSLSPFITGLHGLTHLKL 451
DB 348 DTFOQLLSRLSLNLAWNKIAIHPNFAFSLPSLIKLDLSNLSLSPFITGLHGLTHLKL 407
QY 452 GNHALQSLISSNFPPELVKPYAYQCAFGVCENAYKINQWNGNSMDDLH---- 507
DB 408 GNHALQSLISSNFPPELVKPYAYQCAFGVCENAYKINQWNGNSMDDLH---- 461
QY 508 ---KQDAGNFAQDE----ROLEDFLLDFEEDLKALHSVQCSPPSPGPKPCEHLLDGLW 560
DB 462 ESSKRLPLGLLARAQENHYDQDLDEQLDEM-EDSKPHPSVQCSPTPTGPKPCEYLFESWGI 520

QY 561 RIGWTTIAVLATCNALVTSTVFR-SPLYISPIKILIGVIAAVNMMLTGVSSAVLAGVAF 619
DB 521 RLAVMAIYLLSGLNGLVLLTFVAGGPVLPVPEVFGAIGANTLTGISCGLLASVDAL 580
QY 620 TFGSFARHAGWENGVCVHIGLSIPASBSSEVLLTLAALERGFSVKYSAKFETKAPFS 679
DB 581 TFGSFARHAGWENGVCVHIGLSIPASBSSEVLLTLAALERGFSVKYSAKFETKAPFS 640
QY 680 SLKVIILLCALLTMAAVPLLLGGSKYGASPLCLPL--PFGPESTMGVMVALILNLSLCF 737
DB 641 SVRAGVLGCLALAGLAAALPLASVGEYCASPLCLPYAPPEGQPAALGFTVALVMNSFCF 700
QY 738 LMWTIATYKLYCNLDKGLDENIWCMSVKHIALLLFTNCILNCVPVAFISFSSLIHLTFIS 797
DB 701 LVAGAYIKLYCDLPRGDFEAVWDCAMVRHVAWILFADGLLYCVPVAFISFASMLGFPVT 760
QY 798 PEVTKFILLVVVPLPACLNPLLYLILFNPHFEDLVSLRKQTVWTRSKHPSLSMNSDDV 857
DB 761 PEAVSVLLVVLVPLPACLNPLLYLILFNPHFEDLVSLRKQTVWTRSKHPSLSMNSDDV 815
QY 858 EKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVVTESCHLSSVAFVPC 906
DB 816 EKQSCDSTQALVAFSDVDLILEASEAGRP----PGLETYGFPSVTLISC 860
RESULT 5
QYUY15
ID Q6UY15 PRELIMINARY; PRT; 915 AA.
AC Q6UY15
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Gonadotropin receptor.
GN ORFNames=UNQ6427;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Sehagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL: AY358119; AAQ88486.1; -;
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO: GO:0004872; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR002131; Gphrmn_receptor.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003885; LRR_cyst.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam: PF00560; LRR_1; 15.
DR PRINTS: PR00373; GLYCHORMONER.
DR PRINTS: PR00237; GPCR RHODOPSIN.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00365; LRR_SD22; 5.
DR SMART: SM00369; LRR_TYP; 14.
DR Receptor.
KW RECEPTOR
SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DBB555F4 CRC64;

FT REPEAT 55 79 LRR 1.
 FT REPEAT 81 103 LRR 2.
 FT REPEAT 104 127 LRR 3.
 FT REPEAT 128 151 LRR 4.
 FT REPEAT 153 175 LRR 5.
 FT REPEAT 176 199 LRR 6.
 FT REPEAT 201 223 LRR 7.
 FT REPEAT 224 247 LRR 8.
 FT REPEAT 248 270 LRR 9.
 FT REPEAT 272 294 LRR 10.
 FT REPEAT 318 341 LRR 11.
 FT REPEAT 342 366 LRR 12.
 FT REPEAT 368 387 LRR 13.
 FT REPEAT 388 411 LRR 14.
 FT REPEAT 413 435 LRR 15.
 FT DISULFID 618 693 By similarity.
 FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 292 292 F -> S (in Ref. 1; AAF68989).
 FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
 FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
 SQ SEQUENCE 951 AA; 104460 MW; 550C2DFC22CA1BB CRC64;

Query Match 44.5%; Score 2091.5; DB 1; Length 951;
 Best Local Similarity 50.7%; Pred. No. 2.9e-116;
 Matches 430; Conservative 124; Mismatches 271; Indels 23; Gaps 6;

QY 10 LSLPVLQATGSSPRGVLRRGPTCHCHCPDGMRLLRVDCSDLGSELSPNLSVFTS 69
 DB 5 LGLLCFLALGLGSGAPPLCAAPCSDGD----RRVDCSGKGLTAVPEGLSAFTQ 60

QY 70 YLDLSMNNISQLPNPLSLRLEELRLAGNALTYIPKGAFTGLSLKVLMLQNNOLRHV 129
 DB 61 ALDISMNNITQPEDAFKNPPLEELQAGNDLSFIHPKALSGKLKELVLTQNNQLKTV 120

QY 130 PTEALQNLRSQSLRDANDHISYVPPSCFSGLSLRLHMLDNDALTEIPVQAFRSLSALQ 189
 DB 121 PSEAIRGLSALQSLRDANDHITSVPEDSPEGLVQLRLHMLDNDLSLTPVHPVLSNLTQ 180

QY 190 AMTLANKIHHIPDYAFGNLSLVHLHNNRIHSLGKCPDGLHSLETLDLNNYNDLEF 249
 DB 181 ALTLANKISSIPDPAFTNLSSVLVHLHNNKIRGLSQHCPDGLDNLLETLDLSYNNLGEF 240

QY 250 PTAITLNLKELGPHSNIRISPEKAFVGNPSLTITHFYNDPIQFVGRSAFQHLPELRT 309
 DB 241 PQAIRKARPSLKELGPHSNISVIPDGFNDGNPLRTIHLNDPLSFVGNAPFNLSDLHS 300

QY 310 LTLNGASQITPEPDLTGTANLESILTGAQISSLPQTVCNQLPNLQVLDLSYNNLEDLPS 369
 DB 301 LVIRGASVQQFPNLTGTVHLESILTGTAKISSIPNNLCQEQKMLRLDLSYNNIRDLPS 360

QY 370 FSVCKLOKIDLRHNEIYEKVDVTQQLLSRLSLANLWKNKIAIHPNAPSTLPSLIKLDL 429
 DB 361 FNGCHALEEISIQRNQIYQIKETGTFQGLSLRLDLRSRLIHEIHSRAFATIGPTINLDV 420

QY 430 SSNLLSPFTGLHGLTHLKLGTGNHALQSLSENPELKVTEMYPAYOCCAFGCENAY 489
 DB 421 SFNELTSPTEGLNGLNKLQVGNFKLEALAKDFVNLRSLSVPYAOCCAFWCDSYA 480

QY 490 KISQNNKGDSSMDLHKQAGMFAQODERDLEDLDFEDLKALHS---VQCSPPG 546
 DB 481 NL-----NTEDSLDHSAVEAKGTADAANVTS---TLENEE-----HSQIIHCTPTG 527

QY 547 PPKCEHLLDGMLIRIGVWTHIAVLATCNALVTSTVFRSPYISPIKLLIGVIAANNMLT 606
 DB 528 AFKPCVYLLGSMWIRLTVFPLVALFFNLVILATFASCTSLPSSKLFGLIGLSVSNLFW 587

QY 607 GVSSAVLAGVDAFTGSGFARHCAGWGVGCHVIGFLSIFASESSVFLTLAALRGFSV 666
 DB 588 GIYTGILTFLDVAGWGRFAEFGIMWETGSGCKVAGFLAVFSSESAIFLLMLATVRSLSA 647

QY 667 KYSAKFETKAPFSSSLKVIILLCALLALTMAAVPLPLGGSKYGIASPLCLPLPFGPSPMTGMY 726
 DB 648 KOIMKNGKSNHLKQPRVAALLAFLATGATVAGCPFLPHRGYSASPLCLPPTGPTGTSLGFT 707

QY 727 VALIILNSICELMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNCPVAFUS 786
 DB 708 VTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSSENSOSMIKHWLVAVLIFTNCIFPCPVAFVS 767

QY 787 FSSILNLTPISEVIKFTLLVVVPLPACLNPLLYILFNPHFKEDLSVLKQTYVWTRSKH 846
 DB 768 FAPLITATISPEIMKSVTLFPPLPACLNPLVYVFPFKPKEDWKLLKRRV---TKKSG 824

QY 847 FSLMSINS 854
 DB 825 SVSVSISS 832

RESULT 7
 LGR4 RAT
 ID LGR4 RAT STANDARD; PRT; 951 AA.
 AC Q922H4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
 GN Name=Gpr48; Synonyms=Lgr4;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA MEDLINE=93065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
 RX Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
 RT "Characterization of two lgr genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";
 RL Mol. Endocrinol. 12:1830-1845 (1998).
 CC -!- FUNCTION: Orphan receptor.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL; AF061443; AAC77910.1; --
 DB HSSP; Q9BZR6; IOZN.
 DB RGD; 628615; Gpr48.
 DB InterPro; IPR000276; GPCR_Rhodpsn.
 DB InterPro; IPR002131; Gqhrmn_receptor.
 DB InterPro; IPR001611; LRR.
 DB InterPro; IPR000372; LRR_Nterm.
 DB InterPro; IPR003591; LRR_typ.
 DB Pfam; PF00001; 7tm_1; 1.
 DB Pfam; PF00560; LRR; 15.
 DB Pfam; PF01462; LRRNT; 1.
 DB PRINTS; PR00373; GLYCHORMONER.
 DB PRINTS; PR00237; GPCR_HODOPSIN.
 DB PRINTS; PR00019; LEURICHRPT.
 DB SMART; SM00013; LRRNT; 1.
 DB SMART; SM00369; LRR_TYP; 5.
 DB PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
 DB PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat; Signal; Transmembrane.

FT	SIGNAL	1	24	Potential.
FT	CHAIN	25	951	Leucine-rich repeat-containing G protein-coupled receptor 4.
FT				Extracellular (Potential).
FT	DOMAIN	25	544	1 (Potential).
FT	TRANSMEM	545	565	2 (Potential).
FT	DOMAIN	566	575	Cytoplasmic (Potential).
FT	TRANSMEM	576	596	2 (Potential).
FT	DOMAIN	597	619	Extracellular (Potential).
FT	TRANSMEM	620	640	3 (Potential).
FT	DOMAIN	641	661	Cytoplasmic (Potential).
FT	TRANSMEM	662	682	4 (Potential).
FT	DOMAIN	683	703	Extracellular (Potential).
FT	TRANSMEM	704	724	5 (Potential).
FT	DOMAIN	725	756	Cytoplasmic (Potential).
FT	TRANSMEM	757	777	6 (Potential).
FT	DOMAIN	778	783	Extracellular (Potential).
FT	TRANSMEM	784	804	7 (Potential).
FT	DOMAIN	805	951	Cytoplasmic (Potential).
FT	REPEAT	55	79	LRR 1.
FT	REPEAT	81	103	LRR 2.
FT	REPEAT	104	127	LRR 3.
FT	REPEAT	128	151	LRR 4.
FT	REPEAT	153	175	LRR 5.
FT	REPEAT	176	199	LRR 6.
FT	REPEAT	200	223	LRR 7.
FT	REPEAT	225	247	LRR 8.
FT	REPEAT	248	270	LRR 9.
FT	REPEAT	272	294	LRR 10.
FT	REPEAT	318	341	LRR 11.
FT	REPEAT	342	366	LRR 12.
FT	REPEAT	368	387	LRR 13.
FT	REPEAT	388	411	LRR 14.
FT	REPEAT	413	435	LRR 15.
FT	DISULFID	618	693	By similarity.
FT	CARBOHYD	68	68	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	188	188	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	199	199	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	294	294	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	314	314	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	505	505	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	951 AA;	104138 MW; EDD56AC072123461 CRC64;	
Query Match				
Best Local Similarity 44.4%; Score 2088.5; DB 1; Length 951;				
Matches 438; Conservative 118; Mismatches 271; Indels 19; Gaps 7;				
Qy	10	LSLPVLLQATGSSPRGVLLRGCPHCHCEPDGRMLLRVDCSDLGSELPSNLSVPTS	69	
Db	5	LGLLCFLALGLSAGSPGAAPLCAAPCSGDG----	RRVDCSGKGLTAVPEGLSAFTQ	60
Qy	70	YLDLSMNNISQLLPNPLSPFLBELRLAGNALTYIPKGAFTGLYSKLVLMQNNQLRHV	129	
Db	61	ALDISMNNITQLPEDAFKSPFLEELQAGNDLSLIHPKALSGKELKVLTLQNNQLRTV	120	
Qy	130	PTEALQNRSLQSLRDANHTSYPPSCFGLSHLRLWLDNALTETPQAFPSLSALQ	189	
Db	121	PSEAIHGLSALQSLRDANHTSPEDSFGLVQLRHLDNLSLTVTPVPLSNLTLQ	180	
Qy	190	AMTLAKNHHIPYAFGNLSLVVLHLLNNRIHSLGKKCFDGLHSLFTLDLNNYLDEF	249	
Db	181	ALTALANNISIPDFAFNTLSVLVHLNNKIKLSQHCDFGLDNLETLDLNNYLDEF	240	
Qy	250	PTAIRTLSNKEIGFHSNNIRISIEKAFVGNPSLTIHFYDNPTQFVCRSAFQHLPELRT	309	
Db	241	FOATKALPSLKEIGFHSNISIVIPDGAFGGNPLRTIHLNDPLSFVGNISAPHLNLSLHC	300	
Qy	310	LTNLGASQITFPDPLTGANLESITLTCQAQSSLPQTCVQCNQLPNLQVLDLSYNLLEDLPS	369	
Db	301	LVIRGASLVQFNNLTGTVHLESITLGTCKLSSIPDDLCCQKMLRTLDLSYNNIRDLPS	360	
Qy	370	FSVCQKQKIDLRNHEIYEKVDFTFQQLLSRLSNLAWNKIAITHPNFSTLPSLIKLDL	429	
Db	361	FNGCRALBEISLQRNQISLIKENTFQGLTSLRIILDSRLNRIEHSQAFKLTITNLIDV	420	

Qy	430	SSNLLSFPFITGLHGLTHLKLGTGNHALQSLISSNPFPELKVIEPMPYAYQCCAFGVCENAY	489	
Db	421	SFNELTSPFTEGLNGNLQKLKLVGNFKLDALARDFANLRSLSVPIAYQCCAFGWCDSYA	480	
Qy	490	KISQNMKGNDSMDDLHKKDAQMFQADBERDLEDFLLDFFEDLKALHSVQCSPPGPFK	549	
Db	481	NL-----NTEDNSPQEHSVTKEG--ATDAANTVSTAENEHSQIIH---CTPSTGAFK	530	
Qy	550	PCHELLDGLWIRIGVWTIAVLTALTCNALVTSTVPSPLYISPIKLLIGVIAAVNMLTGV	609	
Db	531	PCGYLLGSMWIRLTWPFIFVALLFNLLVITVFPASCSSLPASKLFIGLISVSNLLMGTY	590	
Qy	610	SAVLGAVDATTFGSFARHGAWGVCHVIGLFIIPASSSVFLLTLAALERGFVSVKYS	669	
Db	591	TGILTFLDVAVSWGRFAEFGIWEYTGCKVAGSLAVFSSSAVFLTLAAVERSVFAKDL	650	
Qy	670	AKFETKAPFSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGPESTWGMVAL	729	
Db	651	MKGKSSHLRQFQVAALLALLGAAGVCGFPLFHGGQYSASPLCLPFPFTGETPSLGTVTL	710	
Qy	730	ILNSLCFLAMTITATYTKLYCNLDKDLNWDSCNVGHIALLLFTNCILNCPVAFLSFS	789	
Db	711	VLLNSLAFLLMAITTYTKLYCNLEKEDLSNSQSSVVKHVAWLIFTNCIFPCVAFPSFAP	770	
Qy	790	LINLTFTSPVIFKILLVVPPLPACLNPLLYLILFNPFPKEDVLRLQRTYVWTRSKHPSL	849	
Db	771	LITATISPEIMKSVTLIFPPLPACLNPLVYVFNFPKEDKLLKRRV---TR-KHGSV	826	
Qy	850	-MSINS	854	
Db	827	SVSISS	832	
RESULT 8				
ID	Q8N537	PRELIMINARY;	PRT;	927 AA.
AC	Q8N537;			
DT	01-OCT-2002	(TREMBlrel. 22, Created)		
DT	01-OCT-2002	(TREMBlrel. 22, Last sequence update)		
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)		
DE	GR48	protein.		
OS	Homo sapiens	(Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalski U., Smallos D.B., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			

Strausberg R.;
Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; BC033039; AAH33039.1; -;
HSSP; P25147; 1DOB.
GO; GO:0016021; C: integral to membrane; IEA.
GO; GO:0016500; F: protein-hormone receptor activity; IEA.
GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
InterPro; IPR000276; GPCR_Rhodopsn.
InterPro; IPR002131; Gphrmn_receptor.
InterPro; IPR001611; LRR.
InterPro; IPR000372; LRR_Nterm.
InterPro; IPR003591; LRR_Typ.
InterPro; IPR007087; Znf_C2H2.
Pfam; PF00001; 7cm_1; 1.
Pfam; PF01462; LRRNT; 1.
Pfam; PF00560; LRR_1; 14.
PRINTS; PR00373; GLYCHORMONER.
PRINTS; PR00237; GPCR_Rhodopsn.
PRINTS; PR00019; LEURICHRPT.
SMART; SM00013; LRRNT; 1.
SMART; SM00369; LRR_TYP; 5.
PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN 1.
SQ SEQUENCE 927 AA; 101676 MW; C7B2F0C40E584C58 CRC64;

Query Match 43.1%; Score 2024.5; DB 2; Length 927;
Best Local Similarity 49.2%; Pred. NO. 2.8e-112;
Matches 420; Conservative 125; Mismatches 254; Indels 55; Gaps 9;

QY 6 LGVL--LSLVLQLATGSSPRSGVLRGCPHCHCEPDGRMLLRVDCSDGLSELPSN 63
DB 5 LGLCLFLALGILGAGPGGAPPL-----CAAPCSGDG----RRVDCSGKGTAVPEG 54
QY 64 LSVETSYLDLSMNNTISQLLPNPLSLRPLEELRLAGNALTYIPKGAFTGLSLKVLMLQN 123
DB 55 LSATF-----QALQAGNDLSFIHPKALSGELKVLTLQN 90
QY 124 NQLRHVPTEALQNLRSQSLRDANHSYVPSPSCFSLHSLRHLDNALTETIPVQAFR 183
DB 91 NQLKTVPSGARGLSALQSLRDANHSYVPSPSCFSLHSLRHLDNALTETIPVQAFR 150
QY 184 SLSALQMTALNKHITPDYAFGNLSLAVLHNNRIHSLGKCFGLHSLTLDNY 243
DB 151 NLPTLQALTLANKISSIPDFAFTNLSSVLHLLHNNKIRSLSCFPGDGLNLETLDNY 210
QY 244 NNLDEFPTAIRTLNKLGLFHSNIRSIPEKAFVGNPSLITIFPDNPIDFVGRSAFOH 303
DB 211 NNLGEFFQAIKALPSLKLGLFHSNISVIPDGFNGDNLTLIHYDNLPLSFVGNSAFHN 270
QY 304 LPELRLTLNGASQITEPDLTGTLNLSLTGTGAQISSLPQTVCNQLPNIQVLDLSYNL 363
DB 271 LSDLSLVIRGASVMVQFPNLGTGVHLSLTGTGKISSIPNLCQEQKMLRTLDLSYNN 330
QY 364 LEDLPSVSCVQLQKIDLRHNEIYEIKVDFTQQLLSRLSMLNKNKIAIHPNFASTLPS 423
DB 331 IRDLPSFNGCHALBEISLQRNIQIQKEGTQGLSLTSLRLSLRNLRIHEISRAFATLGP 390
QY 424 LKLDLSNLSSPPIITGLHGLTLKLTGNHALOSLISSENFPELKVTEMVYVQCCAFG 483
DB 391 ITNLDVSENELTSPTTEGLNGNLKLVGNFKLKEALAAKQFVNLRSLSVPIYVQCCAFW 450
QY 484 VCENAYKISQNWKNKDNMSMDLHKLKAGMFPQADDERDLEPLDFFEDLKLALHS---VQ 540
DB 451 GCDSYANL---NTEDNSLDHSAVEAGKGTADAANVTG---TLNEEE---HSQIIH 497
QY 541 CSPSPGPKPEHLIDGHLIRIGVWTIAVLATCNALVTSTVFSPLYSIPKLLIGVIA 600
DB 498 CTPSTGAFKPCYELGLSGMWRLTVMFIFLVALFNLLVILTFASCTSLPSSEKLFGLIS 557
QY 601 AVNMLTGVSVAVLGADVAFTRGSPARHCAWENGVCVIGFPLSTFASSESVFLLTAAAL 660
DB 558 VSNLPMGIYTGILTFDLDAVSGRFAEFGIWWETGSGCKVAGFLAVFSSSAIFLLMLATV 617

QY 661 ERGFSVKYSAKFETKAPFSSLSKVILLCALLALTAALTAAPVLLGSGSKYASPLCLPFPGE 720
DB 618 ERLSLAKDIMKNGKSNHLKQFRVAALFLAGATVAGCFPIFHRGYSASPLCLPFPPTGET 677
QY 721 STMGVMVALIILNSLCFLMMTITATKLYCNLDKGLNIDWCSVMKXHLALLFTTCILNC 780
DB 678 PSLGFTVTLVLSLAFLLMAVIYTKLYCNLEKEDLSNSQSSMKXKXHLALLFTTCIFPC 737
QY 781 PVAFLSFSLNLTFTISPEVIFKIFLLVVVPLACINPLLYILFNPHFKEDLSLRKQTVY 840
DB 738 PVAFSPAPLITATISPEIMSKSVTLIFFPLACINPLVYVFNPKFKEDKWLKLRV-- 795
QY 841 WTRSKHPSLMSINS 854
DB 796 -TKKSGSVSVSIS 808

RESULT 9
LGR6_HUMAN
ID LGR6_HUMAN STANDARD; PRT; 828 AA.
AC Q9HXB8; Q96K69;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6.
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549; DOI=10.1210/me.14.8.1257;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duijn M., Heueh A.J.W.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signaling mechanism for LGR7.";
RL Mol. Endocrinol. 14:1257-1271(2000).
RN [2]
RP SEQUENCE OF 406-828 FROM N.A.
RX MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;
RA Takeda S., Kadowaki S., Haga T., Takeuchi H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL FEBS Lett. 520:97-101(2002).
RN [3]
RP SEQUENCE OF 431-828 FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yanazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horita T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togai S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshioka Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Taehiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

716 FFPLPACLNVLVYFENPKFKDDWKLKRRV---TR-KHGSVSVSISS 759

Db

RESULT 11

Q8C8A7 PRELIMINARY; PRT; 363 AA.

AC Q8C8A7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched

DE library, clone:C130018C02 product:G protein-coupled receptor 49, full

DE insert sequence.

CN Name=Gpr49;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 403:685-690(2001).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC The FANTOM Consortium;

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

[4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

[5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

[6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RC STRAIN=C57BL/6J; TISSUE=Head;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraka T., Hirozane T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

RA Katoh H., Kawai J., Kohli Y., Kondo S., Konno H., Kouda M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,

RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,

RA Tagawa A., Takahashi P., Takaku-Akaira S., Takada Y., Tanaka T.,

RA Tomaru A., Toya I., Yasunishi A., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK047873; BAC33180.1; -.

DR HSSP; Q9BZR6; 1OZN.

DR MGD; MGI-1341817; Gpr49.

DR GO; GO:0005615; C:extracellular space; TAS.

DR GO; GO:0016021; C:integral to membrane; TAS.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000372; LRR_Nterm.

DR InterPro; IPR003591; LRR_typ.

DR Pfam; PF01462; LRRNT; 1.

DR Pfam; PF00560; LRR; 1; 9.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00013; LRRNT; 1.

DR SMART; SM00369; LRR_TYP; 6.

KW Receptor.

SQ SEQUENCE 363 AA; 39879 MW; D5E2FC4449FCE2CO CRC64;

Query Match 34.2%; Score 1608; DB 2; Length 363;

Best Local Similarity 88.1%; Pred. No. 6.1e-88;

Matches 317; Conservative 14; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MDTSLGLVLLSLPVLQLATGSGSPRSGLVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Db 1 MDTSCVENLLSLALLQLVAAGSGPDAPRGCFSCHELDGRMLLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYLDLSMNNISQLPNPLPSLRFLERLRLAGNALTYIPKGAFTGLYSKVL 120

Db 61 PSNLSVFTSYLDLSMNNISQLPASLLHRLCFLEXLRRLAGNALTYIPKGAFTGLYSKVL 120

Qy 121 LQNNLRHVPTALQNLRLSLQSLRLDANHSYVPPSCFSLHSLRHLWDDNALTIPVQ 180

Db 121 LQNNLRHVPTALQNLRLSLQSLRLDANHSYVPPSCFSLHSLRHLWDDNALTIPVQ 180

Qy 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRHSGLKCKPDLHSLTLD 240

Db 181 AFRSLALQAMTLALNKIHHIADYAFGNLSLVVHLHNNRHSGLKCKPDLHSLTLD 240

Qy 241 LNYNNLDEPPTAINTLNLKELGFSNNIRSIKPAFVGNPSLTIHFYDNPIDFVGRSA 300

Db 241 LNYNNLDEPPTAINTLNLKELGFSNNIRSIKPAFVGNPSLTIHFYDNPIDFVGRSA 300

Qy 301 FOHLPELRTITLNCASOITFPPDLTGTALESLLTCAQISSLPQTVCNOLPNLOVLDS 360

Db 301 FOHLPELRTITLNCASOITFPPDLTGTALESLLTCAQISSLPQTVCNOLPNLOVLDS 360

RESULT 12

Q6HA06 PRELIMINARY; PRT; 1093 AA.

AC Q6HA06;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Glycoprotein hormone receptor.

CN Name=gpr;

OS Crassostrea gigas (Pacific oyster).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;

OC Ostreoida; Ostreidae; Crassostrea.

OX NCBI_TaxID=29159;

RN [1]

RP SEQUENCE FROM N.A.

RA Herpin A., Favrel P.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

DR EMBL; AJ549813; CAD71143.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.


```
Db 274 RRNOISEITSGALTNLRLKVLVDNDSSLSMPVGLNMLLMQLEISANNRIRVWSKGF 333
Qy 278 VGNPSLTIHYDNPQIFVGRSAQHLPRLTLTLNGASQITTEPDLTGTTANLESLLTG 337
Db 334 PKN--LVSLLDKSNPLAGIKFAGLQNMPLRKLJLSDVRGVLNPLPDLGCTSLVRLDR 391
Qy 338 AQISLSPQVQCNQNLQVLDLNLLDLPSFVQCQKQID----- 380
Db 392 ANLSKIPDHICKTSPRLSLDKSNLLSIINVTNCRDLRLDLASNRIISLHGAPSSL 451
Qy 381 -----LRNHEIYEIKVDTFQQLSLRSNLNLANWKIAIHPNAPSTLPISLIKDLSSNL 433
Db 452 QQLHDLNLSNNEIESIPHDAFVGLVRLQVLDMSNRVFFIHADAFPLKLEDLNGLNL 511
Qy 434 LSSPFTGLHGLTHLKTGNHALQSLISSENFPELKVIEPMPYAVQCCA----- 481
Db 512 FPQLPTAGLELLHLKTFNNPHLREFPEPPRIQTLVLVSAYHCCSFPLTTALPKVP 571
Qy 482 --FGVCENA-YKISNQ-----WNKGDSN---SMDDLHKK-----DAGMFOA 516
Db 572 NTFIRENVLPPTNEPDMSLWNSYNDIWPQLQNLKSKFGTQINDLLNAYGAEVGSYPS 631
Qy 517 QDERDLEDFDLDFEDLKALH-----SVQSPSPGPFKPCBHLLDGMLIRIGWTTIA 568
Db 632 GHVPTFPDEY--FEDELGITHASPTAOPGSIQCLPEPGPFLPCQDLFDWMTLRCGVVVF 689
Qy 569 VLALTCNALVTSTVFRSPLYISPIKLIGVITAAVNMULTGVSSAVLAGVDAFTFSPARHG 628
Db 690 LLAMLGNTVVVFLIFSRKMDVPRFLVCNLAARADFNGIYGLVAVVDASTLGEFRMYA 749
Qy 629 AWWENGVCVIGFISLIPASBSSVFLTLTALALRGFVSKYSAKPEKAPSSSLKVIILL 688
Db 750 IPWQMSAGCKLSGLFVAVLSLSVYTLAVTLERNYALTHMLNKLRLSHASYIMTVG 809
Qy 689 ALLALTAAPVLLGSKYKYGASPLCLPLFGE-PTSMGYMVALILNLSLCFLMWTIATKL 747
Db 810 WTFATVAVLPLGVDYRVPAVCLPFEIQGTGSLAYVVFMPFINGVAFILMGCYLKM 869
Qy 748 YCNLDKGDLENWDCSMVKHTALLFTNCILNCVAPLSPSLNLTAFISPEVIFKILV 807
Db 870 YCAIRGQAMNSDRIAKRNALLVFTDFICWSPIAFSLTAVFGLHLISLEQAKVFTVF 929
Qy 808 VVPLPACILNPLIYLFNPFKEDIVSLRKQ-----TVVTRSRKHPSLMSINSDDVEKQS 861
Db 930 ILPLNSCCNPLYALLTKQFKKQKCVLICKATEESRVTRIGCRHSNFS-----NRHT 983
Qy 862 CDSTOALVFTSSITVDLPP 882
Db 984 PANTNSLVERSCK---ELPP 1000

RESULT 14
Q7KTAO PRELIMINARY; PRT; 1360 AA.
AC Q7KTAO;
AT Q7KTAO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE CG8930-PA (CG8930-pb).
GN Name=xk; ORFNames=CG8930;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_taxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA
```

```
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabriellian A., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostali D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou X., Smith H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Weinstock G., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Svirskas R., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RA melanogaster euchromatic genome sequence."
RA Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RA "The transposable elements of the Drosophila melanogaster euchromatin:
RA a genomic perspective."
RA Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.B., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RA "Annotation of the Drosophila melanogaster euchromatic genome: a
RA systematic review."
RA Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (SEP-2002) to the ENBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
```


Db 181 VVLDLGNNNLTKEANSFFMAPNLEDITLSDNSIINMDPNAFYGLAKLKLKLSLQNCGLK 240
Qy 128 HVPTEALQNRSLQSLRLDANHSYVPPSCFSGHSLRHLWLDNALTEIPVQAFRSLSA 187
Db 241 SLPPSQFQGLAQTLQSLQNGNALVSLQDCLGHQKRLTLRLEGNLFYRIPTNALAGLRT 300
Qy 188 LOAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSIGKCCFDGLHSLETLDLANNLD 247
Db 301 LEALNLSNLTITINDEDFPPMPNLI VLLLRNQIMKISAGALKNLTKALVLEDDNLIS 360
Qy 248 EFTAIRTLNKLGFHSGNIRSIPEKAFVGNPSLITIHFDNPIQFVGRSAFQHLPEL 307
Db 361 SLPEGLSKLSQQLSITSNLRWINDELPMSQMLDMR--ANPLSTISAGAFRGM SKL 418
Qy 308 RTLTNGASQITEPPDLTGTANLESFLTGAQISSLPQTVCNQLPNLQVLDLSYNLLEDL 367
Db 419 RKLTLSDVRLTSPPELEACHALEILKLDRAGIOEVPANLCROTTPRLKSLKLTNSLKRI 478
Qy 368 PSFVCCQKQKIDLRHNEIYEIK-----VDTFQQLLSLSL 403
Db 479 PNLSSCRDLRLDSSNQIEKIQGKPNGLKQLNDLLSNRIKALPQDAPQGIKQLLL 538
Qy 404 NLAWNKIAIHPNAFSLPSLIKLDSSNLSFPITGLHGLTHLKLGTGHALQSLISSE 463
Db 539 DLEGNISYIHKEAFSGFTALEDLNIGNNIPPELPESGLRALHLKTFNPKLREFPPD 598
Qy 464 NPPELKVIEPYYQCCAF-----GVCE 486
Db 599 TFPRIOTLILSYAHCCAFPLVAMSSQKTSQVQEAFLPSPDAEFDFWTLMNNSMMNIWP 658
Qy 487 NAYKISNQ-----NNKGDNSMDLHKDAG-----MFOAQDERDLEDLDF--- 529
Db 659 QMHNLSKQLGASMDPNETAINFNEEQLOTGQIATSYNEEYFEEDHVGSGPATGYGFG 718
Qy 530 -----EEDLKALHSVQSPSPGPKCEHLLDGLIRIGVWTIAVLALTCNALVTST 581
Db 719 TGLPSGMSTEDFQPGSVQCLPMGFLPCADLFDWTLRCGVWVFLSLGNGTVFV 777
Qy 582 VFRSPLYISPIKLLIGVIAVNMUTGVSSAVLAGVDAFTPGSPARHGAWENGVGCHVIG 641
Db 778 LLCRSKMDVPRFLVNCNLAADFFMGIVYLGILAVDAATLGEFRMFAIPQMSVLCQLSG 837
Qy 642 FLSTIFASESSVFLTLAALERGSVKYSKAFETKAPFSSSLKVIILLCALLATMAAVPLL 701
Db 838 FLAVLSSELSVYTLAVITLERNYAITRAIHUNKLSLKQAGYINSGVWFALINMLPLV 897
Qy 702 GGSYKYGASPLCLPFGGE-PSTMGYVVALILNLSLCFLMMTIAYTKLYCNLDKGDLENIW 760
Db 898 GVS DYRKFAVCLPFTTTGPASTYVISLMEFINGCAFLTLMGCYLKMYWAIRGSQAWTN 957
Qy 761 DCSMVKHIALLFTNCILNCPVAFLSFSSLINLTFISPEVIKFI LLVVVLPACLNPLY 820
Db 958 DSRIAKRMALLVFTDFLCWSPIAFFSITAFGLQLISLEQAKIFTVFVPLNLSCCNPFLY 1017
Qy 821 ILFNPHEKEDLVSLRKQTYVWTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITYDL 880
Db 1018 AIMTKQFKKOCVIL-----CKH-----FEESRVVGGGCGRGA VARTKEG---DL 1060
Qy 881 PPSSVPSPAYPVTESC 896
Db 1061 PPLLPAAVAHPPGC 1076

Search completed: July 12, 2005, 07:36:34
Job time : 115 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:32 ; Search time 614 Seconds
(without alignments)
570.689 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGLVLLSLPVLQLAT.....PAYPTESCHLSLVAFPVCL 907

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726218 seqs, 386331768 residues

Total number of hits satisfying chosen parameters: 1726218

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	10 US-09-965-536A-16	Sequence 16, Appl
2	4702	100.0	907	10 US-09-970-944-26	Sequence 26, Appl
3	4702	100.0	907	14 US-10-251-385-264	Sequence 264, App
4	4702	100.0	907	14 US-10-225-567A-422	Sequence 422, App
5	4702	100.0	907	15 US-10-295-027-849	Sequence 849, App
6	4702	100.0	907	15 US-10-295-027-946	Sequence 946, App
7	4702	100.0	907	15 US-10-295-027-1331	Sequence 1331, Ap
8	4702	100.0	907	16 US-10-751-736-84	Sequence 84, Appl
9	4702	100.0	907	17 US-10-482-029-158	Sequence 158, App
10	4701	100.0	1145	17 US-10-505-486-103	Sequence 103, App
11	4696	99.9	907	14 US-10-251-385-278	Sequence 278, App

Query Match 100.0%; Score 4702; DB 10; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

RESULT 1

US-09-965-536A-16
; Sequence 16, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMYS,
; FILE REFERENCE: DOO41NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/310,436
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1.
; SEQ ID NO 16
; LENGTH: 907
; TYPE: PRT
; ORGANISM: HUMAN
US-09-965-536A-16

12	4591	92.8	907	10	US-09-970-944-27	Sequence 27, Appl
13	4586	96.9	883	15	US-10-295-027-484	Sequence 484, App
14	4586	96.9	883	15	US-10-295-027-1330	Sequence 1330, Ap
15	4586	96.9	883	15	US-10-173-999-28	Sequence 28, Appl
16	4048	86.1	907	14	US-10-271-078-10	Sequence 10, Appl
17	3598	76.5	693	17	US-10-851-470-4	Sequence 4, Appl
18	2534	53.9	967	10	US-03-851-595-2	Sequence 2, Appl
19	2534	53.9	967	15	US-10-684-667-2	Sequence 2, Appl
20	2527.5	53.8	1205	17	US-10-505-486-64	Sequence 64, Appl
21	2525.5	53.7	967	10	US-09-851-595-11	Sequence 11, Appl
22	2525.5	53.7	967	15	US-10-331-496A-95	Sequence 95, Appl
23	2525.5	53.7	967	15	US-10-664-667-11	Sequence 11, Appl
24	2525.5	53.7	967	16	US-10-737-450-32	Sequence 32, Appl
25	2525.5	53.7	1005	14	US-10-176-847-90	Sequence 23, Appl
26	2492.5	53.0	928	10	US-09-970-944-23	Sequence 23, Appl
27	2465	52.4	940	10	US-09-970-944-6	Sequence 6, Appl
28	2413	51.3	915	15	US-10-295-027-1336	Sequence 1336, Ap
29	2413	51.3	915	16	US-10-783-528-101	Sequence 101, App
30	2410	51.3	915	15	US-10-331-496A-68	Sequence 68, Appl
31	2409	51.2	915	14	US-10-270-336-5	Sequence 5, Appl
32	2389	50.8	893	10	US-09-970-944-34	Sequence 24, Appl
33	2246	47.8	948	16	US-10-398-036-10	Sequence 10, Appl
34	2106.5	44.7	951	17	US-10-935-190-5	Sequence 5, Appl
35	2103.5	44.7	1189	17	US-10-505-486-126	Sequence 126, App
36	2094.5	44.5	951	17	US-10-851-470-2	Sequence 2, Appl
37	2088.5	44.4	951	14	US-10-276-340-3	Sequence 3, Appl
38	2070.5	44.0	951	14	US-10-225-567A-528	Sequence 528, App
39	1980	42.1	823	14	US-10-270-336-6	Sequence 6, Appl
40	1980	42.1	828	10	US-09-970-944-25	Sequence 25, Appl
41	1980	42.1	828	14	US-10-225-567A-581	Sequence 581, App
42	1374	42.0	794	14	US-10-270-336-7	Sequence 7, Appl
43	1373	42.0	828	14	US-10-270-336-2	Sequence 2, Appl
44	1844	39.2	736	10	US-09-851-595-8	Sequence 8, Appl
45	1844	39.2	736	15	US-10-664-667-8	Sequence 8, Appl

```
QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVYPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVYPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSGLGKCFGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSGLGKCFGLHSLETL 240
QY 241 LNNYNNLDEFFTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPV 300
Db 241 LNNYNNLDEFFTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPV 300
QY 301 FOHLPELRTLNGASQITTEPDLTGTTANLESLLTGAQISSLPQTVCNQLPNQVL 360
Db 301 FOHLPELRTLNGASQITTEPDLTGTTANLESLLTGAQISSLPQTVCNQLPNQVL 360
QY 361 YNLLLEDLPSEFVCOKLOKIDLRHNEIYEIKVDYTFQOLLSLRSLNLANWKIAIHPNAPST 420
Db 361 YNLLLEDLPSEFVCOKLOKIDLRHNEIYEIKVDYTFQOLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLISSENPELKVIEMPVAYOCC 480
Db 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLISSENPELKVIEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDDAGMFOAODERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDDAGMFOAODERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSFARHGAWWENGVCVHIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLAGVDAFTFGSFARHGAWWENGVCVHIGFLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPFGEP 720
QY 721 STMGYVVALIILNSLCFLMWTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMWTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV 840
Db 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTECHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTECHLSS 900
QY 901 VAFVPC 907
Db 901 VAFVPC 907
```

RESULT 2

US-09-970-944-26
; Sequence 26, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkete, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-26

Query Match 100.0%; Score 4702; DB 10; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVYPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVYPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSGLGKCFGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSGLGKCFGLHSLETL 240
QY 241 LNNYNNLDEFFTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPV 300
Db 241 LNNYNNLDEFFTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPV 300
QY 301 FOHLPELRTLNGASQITTEPDLTGTTANLESLLTGAQISSLPQTVCNQLPNQVL 360
Db 301 FOHLPELRTLNGASQITTEPDLTGTTANLESLLTGAQISSLPQTVCNQLPNQVL 360
QY 361 YNLLLEDLPSEFVCOKLOKIDLRHNEIYEIKVDYTFQOLLSLRSLNLANWKIAIHPNAPST 420
Db 361 YNLLLEDLPSEFVCOKLOKIDLRHNEIYEIKVDYTFQOLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLISSENPELKVIEMPVAYOCC 480
Db 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLISSENPELKVIEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDDAGMFOAODERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDDAGMFOAODERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSFARHGAWWENGVCVHIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLAGVDAFTFGSFARHGAWWENGVCVHIGFLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPFGEP 720
QY 721 STMGYVVALIILNSLCFLMWTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMWTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV 840
Db 781 PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV 840
```

Db 781 PVAFSPSSLINLTFISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDIVLSLRKQTVV 840
Qy 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 3
US-10-251-385-264
; Sequence 264, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-264

Query Match 100.0%; Score 4702; DB 14; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGGRMLLRVDCSDGLSEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGGRMLLRVDCSDGLSEL 60

Qy 61 PSNLGVFTSYDLDLNNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKLYM 120
Db 61 PSNLGVFTSYDLDLNNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKLYM 120

Qy 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSLRLHLLDNNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSLRLHLLDNNALTEIPVQ 180

Qy 181 AFRSLSALOQMTLALNKTHIIPDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLSALOQMTLALNKTHIIPDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLETLD 240

Qy 241 LNYNNLDFPTAIRTLNSLKGELPHSNNIRSIPEKAFVGNPSLTIHFYDNPVGRSA 300
Db 241 LNYNNLDFPTAIRTLNSLKGELPHSNNIRSIPEKAFVGNPSLTIHFYDNPVGRSA 300

Qy 301 FOHLPELRTLNGASQITTEPDLTGNTANLSLTLTGAISSLPOTVCNQLPNQVLDLS 360
Db 301 FOHLPELRTLNGASQITTEPDLTGNTANLSLTLTGAISSLPOTVCNQLPNQVLDLS 360

Qy 361 YNLEDLPSFVSCVKQKIDLRHNEIYEIKVDYTFQOLLSLRLSNLAWNKIAI IHPNAPST 420
Db 361 YNLEDLPSFVSCVKQKIDLRHNEIYEIKVDYTFQOLLSLRLSNLAWNKIAI IHPNAPST 420

Qy 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480
Db 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480

Qy 481 AFGVCENAYKTSNQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540

Db 481 AFGVCENAYKTSNQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLMIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLMIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGUSSAVLAGVDAFTGSPARHCAMWENGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGUSSAVLAGVDAFTGSPARHCAMWENGCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALMLTMAAVPLLGGSYKASPLCLPLFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALMLTMAAVPLLGGSYKASPLCLPLFGEP 720
Qy 721 STMGYVVALIILNLSLCLFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNLSLCLFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFSPSSLINLTFISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDIVLSLRKQTVV 840
Db 781 PVAFSPSSLINLTFISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDIVLSLRKQTVV 840
Qy 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 4
US-10-225-567A-422
; Sequence 422, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 422
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-422

Query Match 100.0%; Score 4702; DB 14; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGGRMLLRVDCSDGLSEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGGRMLLRVDCSDGLSEL 60

Qy 61 PSNLGVFTSYDLDLNNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKLYM 120
Db 61 PSNLGVFTSYDLDLNNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKLYM 120

Qy 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSLRLHLLDNNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSLRLHLLDNNALTEIPVQ 180

Qy 181 AFRSLSALOQMTLALNKTHIIPDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLSALOQMTLALNKTHIIPDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLETLD 240

QY 241 LNYNLDPEPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
DB 241 LNYNLDPEPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEPPDLTGTTANLESLLTGAQISLLPQTVNCNQLPNIQVLDLS 360
DB 301 FOHLPELRTLTLNGASQITTEPPDLTGTTANLESLLTGAQISLLPQTVNCNQLPNIQVLDLS 360
QY 361 YNLEDLPSFSVCOKLOKIDLRHNEIYEIKVDITFQQLLSRLSLNLANWKIAIHPNAPST 420
DB 361 YNLEDLPSFSVCOKLOKIDLRHNEIYEIKVDITFQQLLSRLSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIAYOCC 480
DB 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOQDERLEDPLDFFEDDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOQDERLEDPLDFFEDDLKALHSVQ 540
QY 541 CSPSPGPFKPECHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPFKPECHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPSSLKVIILLCALLALTAAPVLLGGSKYKASPLCLPLPGEF 720
DB 661 ERGFSVKYSAKFETKAPSSLKVIILLCALLALTAAPVLLGGSKYKASPLCLPLPGEF 720
QY 721 STMGYMWALILLNSLCFLMWTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYMWALILLNSLCFLMWTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFLSLNLATFISPEVIFKILLVVPVLPACLNPLLYTLFNPHEKEDLVSRKQTYV 840
DB 781 PVAFLSFLSLNLATFISPEVIFKILLVVPVLPACLNPLLYTLFNPHEKEDLVSRKQTYV 840
QY 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITYDLPSPSSVPSPAYPVFTSCHLSS 900
DB 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITYDLPSPSSVPSPAYPVFTSCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

RESULT 5

US-10-295-027-849
; Sequence 849, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 849
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-849

Query Match 100.0%; Score 4702; DB 15; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRGLVLLSLPVLQLATGGSPRGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSRGLVLLSLPVLQLATGGSPRGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PSNLVSFTSYLDLSMNNISQLLPNPLPSLRPLESLRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PSNLVSFTSYLDLSMNNISQLLPNPLPSLRPLESLRAGNALTYIPKGAFTGLYSKVL 120
QY 121 LQNNQLRHVPTEALONLSQSLRLDANHSYVPPSCFSGHLSRHLWDDNALTETPVQ 180
DB 121 LQNNQLRHVPTEALONLSQSLRLDANHSYVPPSCFSGHLSRHLWDDNALTETPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240
DB 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240
QY 241 LNYNLDPEPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
DB 241 LNYNLDPEPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEPPDLTGTTANLESLLTGAQISLLPQTVNCNQLPNIQVLDLS 360
DB 301 FOHLPELRTLTLNGASQITTEPPDLTGTTANLESLLTGAQISLLPQTVNCNQLPNIQVLDLS 360
QY 361 YNLEDLPSFSVCOKLOKIDLRHNEIYEIKVDITFQQLLSRLSLNLANWKIAIHPNAPST 420
DB 361 YNLEDLPSFSVCOKLOKIDLRHNEIYEIKVDITFQQLLSRLSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIAYOCC 480
DB 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOQDERLEDPLDFFEDDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOQDERLEDPLDFFEDDLKALHSVQ 540
QY 541 CSPSPGPFKPECHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPFKPECHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660

APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
FILE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1331
LENGTH: 907
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1331

Query Match 100.0%; Score 4702; DB 15; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
Db 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYDLSNMNLSQLLPDLPSIRFLEELRLAGNALTYPKGAFTGLSKVL 120
Db 61 PSNLSVFTSYDLSNMNLSQLLPDLPSIRFLEELRLAGNALTYPKGAFTGLSKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSLRDANHSIVPPSCFSGLSRLHRLWLDNALTPEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSLRDANHSIVPPSCFSGLSRLHRLWLDNALTPEIPVQ 180

Qy 181 AFRSLSALQMTALANKIHHPDYAFGNLSLVLHNNRIHSLGKCKPDLGHSLETLD 240
Db 181 AFRSLSALQMTALANKIHHPDYAFGNLSLVLHNNRIHSLGKCKPDLGHSLETLD 240

Qy 241 LNYNLDLFPFAIRTLNSLKLGHSHNNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA 300
Db 241 LNYNLDLFPFAIRTLNSLKLGHSHNNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA 300

Qy 301 FOHLPELRLTLNGASQITTEPDLTGTANLSLTLTGAQISLSPOTVCNQLPNLQVLDLS 360
Db 301 FOHLPELRLTLNGASQITTEPDLTGTANLSLTLTGAQISLSPOTVCNQLPNLQVLDLS 360

Qy 361 YNLLDLPSFVCQKQKIDLRHNEIYEIKVDYTFQQLSLRSLNLANWKIAIHPNAFST 420

Db 361 YNLLDLPSFVCQKQKIDLRHNEIYEIKVDYTFQQLSLRSLNLANWKIAIHPNAFST 420

Qy 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENFPBLKVLEMPAYQCC 480
Db 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENFPBLKVLEMPAYQCC 480

Qy 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEFLDDFEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEFLDDFEDLKALHSVQ 540

Qy 541 CSPSPGPFKCEHLLDGLWLRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPFKCEHLLDGLWLRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Qy 601 AVNMLTGVSNAVLGADVDAFTFGSPARHGAMWENGCHVIGFLSIFASESVFLLTAAAL 660
Db 601 AVNMLTGVSNAVLGADVDAFTFGSPARHGAMWENGCHVIGFLSIFASESVFLLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLKVILLCALIALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVILLCALIALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720

Qy 721 STMGYVALIILNSLCFLAMMTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLAMMTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780

Qy 781 PVAFLSFSSILNLTFFISPEVIKFTLLVVPVLPACINPLLYLTFNPHKEDLVSRKQTYV 840
Db 781 PVAFLSFSSILNLTFFISPEVIKFTLLVVPVLPACINPLLYLTFNPHKEDLVSRKQTYV 840

Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSSIYDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSSIYDLPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 8

US-10-751-736-84
Sequence 84, Application US/10751736
Publication No. US20040265230A1

GENERAL INFORMATION:

APPLICANT: Wveth

APPLICANT: Martinez, Robert

APPLICANT: Brown, Eugene

APPLICANT: Liu, Wei

TITLE OF INVENTION: CANCERS

FILE REFERENCE: AM100927 (031896-002000)

CURRENT APPLICATION NUMBER: US/10/751,736

CURRENT FILING DATE: 2003-01-06

PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000

PRIOR FILING DATE: 2003-01-06

NUMBER OF SEQ ID NOS: 54873

SOFTWARE: PatentIn version 3.2

SEQ ID NO 84

LENGTH: 907

TYPE: PRT

ORGANISM: Homo sapiens

US-10-751-736-84

Query Match 100.0%; Score 4702; DB 16; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Db 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYDLSNMNLSQLLPDLPSIRFLEELRLAGNALTYPKGAFTGLSKVL 120

Db 61 PSNLSVFTSYLDLNMNIIQSLLPNPLPSLRFLLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNDNALTEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRIHSLGKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRIHSLGKCFDGLHSLETL 240
QY 241 LNNYNNLDEFPPTAIRTLNLSKELGFHSHNNIIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA 300
Db 241 LNNYNNLDEFPPTAIRTLNLSKELGFHSHNNIIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
QY 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSINLAWNKIAIHHNAPST 420
Db 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSINLAWNKIAIHHNAPST 420
QY 421 LPSLKLIDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLISSENPELKVIEPMPAYQCC 480
Db 421 LPSLKLIDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLISSENPELKVIEPMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDQAGFQADDERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDQAGFQADDERDLEDLDFEEDLKALHSVQ 540
QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLCGSKYGASPLCLPLPFGP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLCGSKYGASPLCLPLPFGP 720
QY 721 STMGYMWVALILNLSLCFLMWTIATYKLYCNLDKGDLENIDWCSMWKHIALLFTNCILNC 780
Db 721 STMGYMWVALILNLSLCFLMWTIATYKLYCNLDKGDLENIDWCSMWKHIALLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFTISPEVIRKFIILLVVVPLPACLNPLLYILFNPHFKEDELVSLRKQTV 840
Db 781 PVAFLSFSSLINLTFTISPEVIRKFIILLVVVPLPACLNPLLYILFNPHFKEDELVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 9

US-10-482-029-158
; Sequence 158, Application US/10482029
; Publication No. US20050037445A1

GENERAL INFORMATION:

; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482, 029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 907

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-158

Query Match 100.0%; Score 4702; DB 17; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLQALATGSSPRSGVLLRGCPHCHCEPDGRMLLAVDCSDGLSEL 60
Db 1 MDTSLRGVLLSLPVLQALATGSSPRSGVLLRGCPHCHCEPDGRMLLAVDCSDGLSEL 60
QY 61 PSNLSVFTSYLDLNMNIIQSLLPNPLPSLRFLLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
Db 61 PSNLSVFTSYLDLNMNIIQSLLPNPLPSLRFLLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNDNALTEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRIHSLGKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRIHSLGKCFDGLHSLETL 240
QY 241 LNNYNNLDEFPPTAIRTLNLSKELGFHSHNNIIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA 300
Db 241 LNNYNNLDEFPPTAIRTLNLSKELGFHSHNNIIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
QY 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSINLAWNKIAIHHNAPST 420
Db 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSINLAWNKIAIHHNAPST 420
QY 421 LPSLKLIDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLISSENPELKVIEPMPAYQCC 480
Db 421 LPSLKLIDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLISSENPELKVIEPMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDQAGFQADDERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDQAGFQADDERDLEDLDFEEDLKALHSVQ 540
QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLCGSKYGASPLCLPLPFGP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLCGSKYGASPLCLPLPFGP 720
QY 721 STMGYMWVALILNLSLCFLMWTIATYKLYCNLDKGDLENIDWCSMWKHIALLFTNCILNC 780
Db 721 STMGYMWVALILNLSLCFLMWTIATYKLYCNLDKGDLENIDWCSMWKHIALLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFTISPEVIRKFIILLVVVPLPACLNPLLYILFNPHFKEDELVSLRKQTV 840
Db 781 PVAFLSFSSLINLTFTISPEVIRKFIILLVVVPLPACLNPLLYILFNPHFKEDELVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 10

```
US-10-505-486-103
; Sequence 103, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 103
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Human
US-10-505-486-103

Query Match      100.0%; Score 4701; DB 17; Length 1145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGLSEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGLSEL 60

Qy 61 PSNLSVFTSYLDLSNMNI S QLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120
Db 61 PSNLSVFTSYLDLSNMNI S QLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSVPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSVPPSCFSGLSLRLHLLWLDNNALEIPVQ 180

Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSGLKGCDFGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSGLKGCDFGLHSLETL 240

Qy 241 LNNYNNLDEFFTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300
Db 241 LNNYNNLDEFFTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300

Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCQPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCQPNLQVLDLS 360

Qy 361 YNNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQQLLSLRLSLNLANWKIAIHPNAPST 420
Db 361 YNNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQQLLSLRLSLNLANWKIAIHPNAPST 420

Qy 421 LPSLTKLDSLNNLSLSPFTGLGHGTHLKLTCGNHALQSLISSENPPELVKTEMVAYOCC 480
Db 421 LPSLTKLDSLNNLSLSPFTGLGHGTHLKLTCGNHALQSLISSENPPELVKTEMVAYOCC 480

Qy 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540

Qy 541 CSPSPGPKPCEHLLDGLMIRIGVWTTIAVLAITCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLMIRIGVWTTIAVLAITCNALVTSTVFRSPLYSPIKLLIGVIA 600

Qy 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFAFSESSVELLTAAAL 660
Db 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFAFSESSVELLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
```

```
Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNLSLCFLMMTTIATKLYCNLDKGDLENIWDCSMVKHIALALLFTNCILNC 780
Db 721 STMGYVALIILNLSLCFLMMTTIATKLYCNLDKGDLENIWDCSMVKHIALALLFTNCILNC 780
Qy 781 PVAFLSFSSSLNLTFFISPEVIKFIILLVVVPLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
Db 781 PVAFLSFSSSLNLTFFISPEVIKFIILLVVVPLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTTESCHLSS 900
Qy 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 11
US-10-251-385-278
; Sequence 278, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-278

Query Match      99.9%; Score 4696; DB 14; Length 907;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGLSEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGLSEL 60

Qy 61 PSNLSVFTSYLDLSNMNI S QLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120
Db 61 PSNLSVFTSYLDLSNMNI S QLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSVPPSCFSGLSLRLHLLWLDNNALEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSVPPSCFSGLSLRLHLLWLDNNALEIPVQ 180

Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSGLKGCDFGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSGLKGCDFGLHSLETL 240

Qy 241 LNNYNNLDEFFTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300
Db 241 LNNYNNLDEFFTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300

Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCQPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCQPNLQVLDLS 360

Qy 361 YNNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQQLLSLRLSLNLANWKIAIHPNAPST 420
Db 361 YNNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQQLLSLRLSLNLANWKIAIHPNAPST 420
```

```
Db 361 YNLEDELPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSNPPELKVEMPAYQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSNPPELKVEMPAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Qy 721 STMGYVVALIILNSLCFLMTTAYTKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMTTAYTKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSSLINLTFTISPEVIFILLVVVPLPACLNPLLYIILFNPHEKEDLVSLRKQTVV 840
Db 781 PVAFLSFSSLINLTFTISPEVIFILLVVVPLPACLNPLLYIILFNPHEKEDLVSLRKQTVV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSSITVDLPSSVSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSSITVDLPSSVSPAYPVTESCHLSS 900
Qy 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 12
US-09-970-944-27
; Sequence 27, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; PRIOR FILING DATE: 2002-05-02
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 27
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-27

Query Match 99.8%; Score 4691; DB 10; Length 907;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 905; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60
Qy 61 PSNLSVFTSYLDLSMNISQLLPNLPFLPSLFLFEEELRAGNALTYTPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYLDLSMNISQLLPNLPFLPSLFLFEEELRAGNALTYTPKGAFTGLYSKVLV 120
```

```
Qy 121 LQNNQLRHVPTALQNLRSQSLRLDANHI SYVPPSCFSGLSLRLHLMDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTALQNLRSQSLRLDANHI SYVPPSCFSGLSLRLHLMDDNALTEIPVQ 180
Qy 181 AFRSLSALQAMTLALANKIHHIPDYAFGNLSLVLHLHNNRTHSLGKCKFDGLHSLLETLD 240
Db 181 AFRSLSALQAMTLALANKIHHIPDYAFGNLSLVLHLHNNRTHSLGKCKFDGLHSLLETLD 240
Qy 241 LNNYNNLDFPPTAIRTLNKLPGFHSNNIRSIPEKAFVGNPNLSLITTHFDNPIQFVGSA 300
Db 241 LNNYNNLDFPPTAIRTLNKLPGFHSNNIRSIPEKAFVGNPNLSLITTHFDNPIQFVGSA 300
Qy 301 FOHLPELRLTLNGASQITEFPDLTGTANTLESILTGAQISSLPQTVCNQNLNQLVLDLS 360
Db 301 FOHLPELRLTLNGASQITEFPDLTGTANTLESILTGAQISSLPQTVCNQNLNQLVLDLS 360
Qy 361 YNLEDELPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
Db 361 YNLEDELPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSNPPELKVEMPAYQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSNPPELKVEMPAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Qy 721 STMGYVVALIILNSLCFLMTTAYTKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMTTAYTKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSSLINLTFTISPEVIFILLVVVPLPACLNPLLYIILFNPHEKEDLVSLRKQTVV 840
Db 781 PVAFLSFSSLINLTFTISPEVIFILLVVVPLPACLNPLLYIILFNPHEKEDLVSLRKQTVV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSSITVDLPSSVSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSSITVDLPSSVSPAYPVTESCHLSS 900
Qy 901 VAFVPECL 907
Db 901 VAFVPECL 907
```

```
RESULT 13
US-10-295-027-484
; Sequence 484, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
```

```
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-484

Query Match 96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

Qy 1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCTHCHCEPDGEMLRVDCSDLGSEL 60
Db 1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCTHCHCEPDGEMLRVDCSDLGSEL 60
Qy 61 PSLNSVFTSYLDLSMNISQLPNPLSLRFLLELRAGNALTYPKGAFTGLSLKVL 120
Db 61 PSLNSVFTSYLDLSMNISQLPNPLSLRFLLELRAGNALTYPKGAFTGLSLKVL 120
Qy 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFGLHSRLHMLDNNALTEIPVQ 180
Db 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFGLHSRLHMLDNNALTEIPVQ 180
Qy 181 AFRSLALQMTLALNKIHHIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLETL 240
Db 181 AFRSLALQMTLALNKIHHIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLETL 240
Qy 241 LNYNLDLDEFTAIRTLNLKELGFHNNIRISITPEKAFGNPSLITIFDYDNPQVGRSA 300
Db 241 LNYNLDLDEFTAIRTLNLKEL-----HFYDNPQVGRSA 276
Qy 301 FOHLPELRTLNGASQITEFPDLTGNTANLESRLTGAQISLPTVCNQLPNLQVLDLS 360
Db 277 FOHLPELRTLNGASQITEFPDLTGNTANLESRLTGAQISLPTVCNQLPNLQVLDLS 336
Qy 361 YNLLEDLPSFVCQKQKIDLRHNEIYKVDTFQOILSLRSLNLANWKIAIHPNAEST 420
Db 337 YNLLEDLPSFVCQKQKIDLRHNEIYKVDTFQOILSLRSLNLANWKIAIHPNAEST 396
Qy 421 LPSLTKLDLSNLLSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPYQCC 480
Db 397 LPSLTKLDLSNLLSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPYQCC 456
Qy 481 AFGVCENAYKISNQWNGDGNSSMDLHKDAGMFOAQDERDLEDLFLDFEEDLKALHSVQ 540
Db 457 AFGVCENAYKISNQWNGDGNSSMDLHKDAGMFOAQDERDLEDLFLDFEEDLKALHSVQ 516
```

```
Qy 541 CSPSPGPKPCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 517 CSPSPGPKPCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 576
Qy 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWGVCHVIGFLSIFASESSVFLTLAAL 660
Db 577 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWGVCHVIGFLSIFASESSVFLTLAAL 636
Qy 661 ERGFSVKYSAKFETKAPFSSLVKVIILLCALLATMAAVPLLGSKYGNASPLCLPFGEP 720
Db 637 ERGFSVKYSAKFETKAPFSSLVKVIILLCALLATMAAVPLLGSKYGNASPLCLPFGEP 696
Qy 721 STMGVMVALILLNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 697 STMGVMVALILLNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 756
Qy 781 PVAFLSFSSLINLTFISPEVIKFIILVVVPLPACLNPLLYILFNPHFKEDLVSRKQTV 840
Db 757 PVAFLSFSSLINLTFISPEVIKFIILVVVPLPACLNPLLYILFNPHFKEDLVSRKQTV 816
Qy 841 WTRSKHPSLMSINSDVDVEKQSCDSTQALVTFTSSSITVDLPPSSVPSPAYVPTESCHLSS 900
Db 817 WTRSKHPSLMSINSDVDVEKQSCDSTQALVTFTSSSITVDLPPSSVPSPAYVPTESCHLSS 876
Qy 901 VAFVPC 907
Db 877 VAFVPC 883

RESULT 14
US-10-295-027-1330
; Sequence 1330, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevez, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
```

; SEQ ID NO 1330
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1330

Query Match 96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

QY 61 PNLVSFTSYLDLNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLVSFTSYLDLNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120

QY 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRHLWLDNLTPIVQ 180
DB 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRHLWLDNLTPIVQ 180

QY 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSETLD 240
DB 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSETLD 240

QY 241 LNNNLDDEPPTAIRTLNKLGPHTSNIRSIPEKAFVGNPSLIITHFYDNPFIQFVGRSA 300
DB 241 LNNNLDDEPPTAIRTLNKLGPHTSNIRSIPEKAFVGNPSLIITHFYDNPFIQFVGRSA 276

QY 301 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNLQVLDLS 360
DB 277 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNLQVLDLS 336

QY 361 YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSNLAWNKIATHHNAFST 420
DB 337 YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSNLAWNKIATHHNAFST 396

QY 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVEMPAYQCC 480
DB 397 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVEMPAYQCC 456

QY 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540
DB 457 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ 516

QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 517 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 576

QY 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 660
DB 577 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCVIGFSLIFASESSVFLTLAAL 636

QY 661 ERGFSVKYSAKETKAPPSSLVITLLCALLATWAAVPLLGGSKYGASPLCLPLPFGEP 720
DB 637 ERGFSVKYSAKETKAPPSSLVITLLCALLATWAAVPLLGGSKYGASPLCLPLPFGEP 696

QY 721 STMGYVVALILANSCLFMWTIATKLYCNLDKGLNIWDCSMWKHTALLFTNCILNC 780
DB 697 STMGYVVALILANSCLFMWTIATKLYCNLDKGLNIWDCSMWKHTALLFTNCILNC 756

QY 781 PVAFLSFSSLINLTFTSPEVIFILLVVVPLPACLNPLLIYILFNPHEKEDLVSLRKQTV 840
DB 757 PVAFLSFSSLINLTFTSPEVIFILLVVVPLPACLNPLLIYILFNPHEKEDLVSLRKQTV 816

QY 841 WTRSKHPSLMSINSDVEKQSCDQALVFTTSSITTYDLPPSSVPSPAYPTESCHLSS 900
DB 817 WTRSKHPSLMSINSDVEKQSCDQALVFTTSSITTYDLPPSSVPSPAYPTESCHLSS 876

QY 901 VAFVPC 907
DB 877 VAFVPC 883

RESULT 15

US-10-173-999-28
; Sequence 28, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-28

Query Match 96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

QY 61 PNLVSFTSYLDLNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLVSFTSYLDLNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120

QY 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRHLWLDNLTPIVQ 180
DB 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRHLWLDNLTPIVQ 180

QY 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSETLD 240
DB 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSETLD 240

QY 241 LNNNLDDEPPTAIRTLNKLGPHTSNIRSIPEKAFVGNPSLIITHFYDNPFIQFVGRSA 300
DB 241 LNNNLDDEPPTAIRTLNKLGPHTSNIRSIPEKAFVGNPSLIITHFYDNPFIQFVGRSA 276

QY 301 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNLQVLDLS 360
DB 277 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNLQVLDLS 336

QY 361 YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSNLAWNKIATHHNAFST 420
DB 337 YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSNLAWNKIATHHNAFST 396

QY 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVEMPAYQCC 480
DB 397 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVEMPAYQCC 456

QY 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540
DB 457 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ 516

QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db		517	CSPSPGPKCEHLLDGLHRIIGVWTIAVLATCNALVTSVFRSPLYISPIKLLIGVIA	576
Qy		601	AVNMLTGVSSAVLAGVDFTFGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL	660
Db		577	AVNMLTGVSSAVLAGVDFTFGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL	636
Qy		661	ERGFVKYSAXFETKAPFSSLKVIIILLCALLALTMAAVPLLGSKYGASPLCLPLPGEP	720
Db		637	ERGFVKYSAXFETKAPFSSLKVIIILLCALLALTMAAVPLLGSKYGASPLCLPLPGEP	696
Qy		721	STMGYVALIILNSLCFLMWTIAVTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780
Db		697	STMGYVALIILNSLCFLMWTIAVTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	756
Qy		781	PVAFLSFSSLINLTFISPEVVKFILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	840
Db		757	PVAFLSFSSLINLTFISPEVVKFILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	816
Qy		841	WTRSKHPSLMSINSDDEKQSCDSTQALVFTTSSITTYDLPPSSVPSPAYPVTESCHLSS	900
Db		817	WTRSKHPSLMSINSDDEKQSCDSTQALVFTTSSITTYDLPPSSVPSPAYPVTESCHLSS	876
Qy		901	VAFVPCL 907	
Db		877	VAFVPCL 883	

Search completed: July 12, 2005, 07:48:45
Job time : 621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:51:27 ; Search time 1145 Seconds
(without alignments)
4689.262 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSLGVLSPVLLQLAT.....PAYPTESCHLSVAFVPCPL 907

Scoring table:

OLIGOP 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cpn2.1/USPTO.spool/US10751736/runat.12072005.083640.22162/app.query.fasta.1.1095
-DB=N_Geneseq.16Dec04 -QFMT=fastap -SUPFIX=oli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CCN 1 1.810 @runat.12072005.083640.22162 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPHLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	2724	3 AAA30770	AAA30770 Human G p
2	907	100.0	2724	8 ABZ42816	ABZ42816 Human G p
3	907	100.0	2724	10 ADC22782	ADC22782 Human G p
4	907	100.0	2724	10 ADH14255	ADH14255 Human HG3
5	907	100.0	2724	12 ADI32985	ADI32985 Human G p

6	907	100.0	2724	12	ADO29923	Ado29923 Human GPC
7	907	100.0	2880	11	ADN39796	Adn39796 Cancer/an
8	907	100.0	2880	13	ADQ80249	Adq80249 G protein
9	907	100.0	2973	13	ADR67869	Adr67869 Human HG3
10	907	100.0	3032	11	ADN39627	Adn39627 Cancer/an
11	907	100.0	3032	11	ADN39530	Adn39530 Cancer/an
12	907	100.0	4558	2	AXZ33980	AXZ33980 Human HG3
13	907	100.0	4570	12	ADH12472	Adh12472 Human ste
14	833	91.8	3297	3	AAA30779	AAA30779 Human HG3
15	820	90.4	2724	3	AAA30779	AAA30779 DNA encod
16	820	90.4	2724	10	ADC22796	Adc22796 Human G p
17	820	90.4	2724	10	ADH14269	Adh14269 Mutated h
18	806	88.9	3438	10	ADP70582	Adp70582 Orphan re
19	621	68.5	2851	10	ADB80463	Adb80463 Ovarian c
20	621	68.5	2651	11	ADN39165	Adn39165 Cancer/an
21	621	68.5	2651	11	ADN39795	Adn39795 Cancer/an
22	606	66.8	2082	2	AAZ25344	AAZ25344 Human LGR
23	232	25.6	723	8	ACA04772	ACA04772 cDNA enco
24	202	22.3	606	2	AZA40459	Aza40459 Human gon
25	197	21.7	1790	6	ABQ55077	Abq55077 Human ova
26	51	5.6	2724	12	ADR30213	Adr30213 Mouse GPC
27	51	5.6	3098	13	ADR67870	Adr67870 Mouse HG3
28	23	2.5	2636	5	ABA09697	AbA09697 Human bon
29	19	2.1	60	6	ABN47039	Abn47039 Human spl
30	18	2.0	284	2	AZA40463	Aza40463 Human gon
31	18	2.0	798	3	AAC75792	Aac75792 Human ORF
32	18	2.0	1827	12	ADQ64302	Adq64302 Novel hum
33	18	2.0	2208	6	AAI67925	AAI67925 Human LGR
34	18	2.0	2208	11	ADN02247	Adn02247 Human par
35	18	2.0	2208	12	ADK19412	Adk19412 ORF of pa
36	18	2.0	2376	12	ADO29126	Ado29126 Mouse nov
37	18	2.0	2484	12	ADO29786	Ado29786 Human nov
38	18	2.0	2484	12	ADQ76788	Adq76788 Leucine-r
39	18	2.0	2487	6	ADH50804	Adh50804 Human G-p
40	18	2.0	2711	6	AAI67924	AAI67924 Partial c
41	18	2.0	2711	11	ADN02245	Adn02245 Human par
42	18	2.0	2711	12	ADK19410	Adk19410 Partial c
43	18	2.0	2786	10	ADC16710	Adc16710 Human G-p
44	18	2.0	2901	6	AAI67927	AAI67927 Human LGR
45	18	2.0	2901	6	AAI67921	AAI67921 Mouse LGR

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
XX
AC AAA30770;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.
XX
KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.
XX
OS Homo sapiens.
XX
PN WO200022129-A1.
XX
PD 20-APR-2000.
XX
PF 12-OCT-1999; 99WO-US023938.
XX
PR 13-OCT-1998; 98US-00170496.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Behan DP, Chalmers DT, Liaw CW;
XX
DR WPI; 2000-329165/28.

501 SerSerMetAspAspLeuHisIysIysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db ACAGTATGAGGAGCTTCATAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAGCT 1560
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLeuHisSerValGln 540
Db GACCTTGAGATTTCTGCTTGCCTTTGAGGAGACCTGAAAGCCCTTCATTCAGTGCG 1620
541 CysSerProSerProGlyProPheIysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db TGTTCACCTTCCAGGCCCTTCAACCCCTGTAACACCTGCTTGTGATGGCTGATC 1680
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db AGAATTTGGAGTGTGACCATAGACATGCTGAGCACTTACTTGTAAATGCTTGTGATCTCA 1740
581 ThrValPheArgSerProLeuThrIleLeuSerProIleLeuLeuLeuIleGlyValIleAla 600
Db ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAACCTGTTAATTTGGGGTCAATGCA 1800
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db TTTGGCAGCTTTTGCACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
661 GluArgGlyPheSerValIysTyrSerAlaIysPheGluThrIleAlaPheSerSer 680
Db GAGCGTGGTCTCTGTGAATATCTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC 2040
681 LeuIysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db CTGAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
701 LeuGlyGlySerIysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db CTGGGTGGCAGCAGATGAGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuMet 740
Db ACACCATGGCTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
741 ThrIleAlaTyrThrIysLeuTyrCysAsnLeuLeuAspGlyAspLeuGluAsnIleTrp 760
Db ACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAATATTTGG 2280
761 AspCysSerMetValIysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db GACTGCTCTATGTTAAACACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db CTGTGGCTCTCTCTGTGCT 2340
801 IleIysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db ATTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
821 IleLeuPheAsnProHisPheIysGluAspLeuValSerLeuArgIysGlnThrIleVal 840
Db ATCTTCTTCAATCTCCTTTAGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
841 TrpThrArgSerIysHisIleProSerLeuMetSerIleAsnSerAspAspValGluIysGln 860
Db TGGACAAGATCAAAACACCAAGCTTGTGATCAATTAATCTGATGATGCGAAACAG 2580
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrIleThrAspLeu 880

2581 TCTGTGACTCAACTCAAGCCTTGTAACTTTTACAGCTCCAGATCACTTATGACCTG 2640
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db CTCTCCAGTTCTGTCCTCATCCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
901 ValAlaPheValProCysLeu 907
2701 GTGGCATTTGTCCCATGTCTC 2721

RESULT 2

ABZ42816

ID ABZ42816 standard; DNA; 2724 BP.

XX

AC ABZ42816;

XX

DT 04-MAR-2003 (first entry)

XX

DE Human G protein-coupled receptor GPR49 nucleotide SEQ ID NO:421.

XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.

XX Homo sapiens.

OS

XX WO200261087-A2.

PN

XX 08-AUG-2003.

PD

XX 19-DEC-2001; 2001WO-US050107.

PF

XX 19-DEC-2000; 2000US-0257144P.

PR

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

PA

PI Burner GC, Roush CL, Brown JP;

XX

XX WPI; 2003-046718/04.

DR P-PSDB; ABP81968.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

102e

NG

No specific

Colon Cancer

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute

D	b		181	CCTTCCAAACCTCAGCGTTTCACCTCTCACTAGAACCTCAGTAGAACAACATCATGCTCAG	240
O	y		81	LeuLeuProAsnProLeuLeuProSerLeuArgPheLeuGlulLeuArgLeuAlaGIyAsn	100
D	b		241	CTGTCCGGATTCCCCTGCCAGTCTCGCTTCCTGGAGGAGTTACTGTCTTGCGGGAAC	300
O	y		101	AlalaLeuThrTyrrilleProLySGlyAlaPheThrglYleutyrsrLeulySValLeuMet	120
D	b		301	GCTCTGCATACATAITCCAAGGAGCATTCAGTGGCCCTTTACAGTCTTAAGAATCTTATG	360
O	y		121	LeuGlnAsnIasnGlnLeuArghHisValproThrClualalaleuGlnAsnLeuArgSerLeu	140
D	b		361	CTGCAGAAATAATCAGCTAAGACACGTACCACAGAACTCTGCAGAAATTTGCCAAGCCTT	420
O	y		141	GlnSerLeuArgLeuAspAlaaSnHsiIseserTyrvValProProSerCySPheSerGly	160
D	b		421	CAATCCCTTCGGTCTGGATGCTTAACCAACATACAGTANGTGCCCCACAGCTGTTTCAGTGGC	480
O	y		161	LeuHiSSerLeuArgHisLesLutrpLeuASpAsnAlalaLeuthrGliulleProValGln	180
D	b		481	CTGCATTCCCTAGSCACCTGTGGCTGGATGACAATGCGTTAAcagAAAATcccCGTccAg	540
O	y		181	AlaPheaRgsrLeuSerAlalaLeuGlnAlaMetThrLeualaLeuAnlysIlleHisHis	200
D	b		541	GCTTTTAGAAGTTTTATCGGCATTGCCAGCCATGACCTTGGCCCTGGAAcaAAAAATAcAccCAC	600
O	y		201	IleProaspTyvralaPheGlvAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn	220
D	b		601	ATACcAGAcTAtgCcTtTTgNaacCcTtcCmAgctTtgTaGttTCtPAcATcTCCATTAACAT	660
O	y		221	ArgileHisserLeuGlvLYSLysCYePhEASPGLyleuHisSRLeuGiuthrLeuAsp	240
D	b		661	AGNATCCACTCCCTGGSAAGAAATGCTTIGATGGCTCCACAGCCTAGAGACTTTAGAT	720
O	y		241	LeuaenTyraenAsnLeuApsGluPhEPheProthrAlaIlaargthrLeuSerAsnLeulyS	260
D	b		721	TTAANAATACAAATAACCTTGATGAATTOCCCACTGCAAATTAGGACACTCTCCAACCTTAAA	780
O	y		261	GluleuGlvPheHisSerAsnAsnIleaRGsrIleProGlvlyvsAlaPheValGlvAsn	280
D	b		781	GNACTAGGATTTCTAGCAACAATATCAGGTGCAATACCTGAGAAGCATTTGTAGGCNAc	840
O	y		281	ProSerLeuilleThrlieHisPheTYasPsnProilleGlnPheValGlvArgSerAla	300
D	b		841	CCTTCTCTATTACAATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGACTGCT	900
O	y		301	PheGlnHisLeuProGlvLUeuArghThrLeuthrLeuanGlvAlaSerGlnlleThrGlu	320
D	b		901	TTTTCAACATTTACCTGAActAAGAACACTGCActCTGaATGGTGCTCCAAAAATAACTGAA	960
O	y		321	PheProaspLeuthrGlvThrAlaAsnLeucLuSrLeuthrLeuthrGlvAlaGlnIle	340
D	b		961	TTTCTGTATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACCTGGAGCAcAGATC	1020
O	y		341	SerSerLeuproGlnThrvAlcyASnGlnLeuProAsnLeuinValLeuasPLeusEr	360
D	b		1021	TCATCTCTCTCAACCGTCTGCANTCAGTTACTAATCTCCAAGTGTAGATCTGCT	1080
O	y		361	TyrAsnLeuLeuglvAuspLeuproSerPheSerValCysGlnlySLeuGlnLYsilaeSp	380
D	b		1081	TACAACCTATTAGAAGATTTACCcAGTTTTTTCAGTCTGCCAAAAAGCTTCAGAAAAATTGCAC	1140
O	y		381	LeuarGHISanGlvulletyrGlvulleIeySvalapThrPheGlnGlnLeuLeuSerLeu	400
D	b		1141	CTAAGACATAATGAATACTACGAANAATTAAAGTGTGACACTTTCCAGCAGTGTGCTTAGCCCTC	1200
O	y		401	ARGSerLeuAsnLeuAlatripAsnLYslleAlatlleHisProAsnAlaPheSerThr	420
D	b		1201	CGATCGCTGANNTTGGCTTGGAACAAATTCGTATTATTACCCCAANTGCATTTTCCACT	1260
O	y		421	LeuproSerLeuilleLYseuApLSusSerSerAsnLeuLeuSerPheProfileThr	440
D	b		1261	TTGCATCCCTTAATAAGCTGGACCTATCGCCCAACTCTGTGCTCTTTTCTATAACT	1320

Qy	441	GlyLeuHisGlyLeuThrHisLeuLysLeuLeuThrGlyAsnHisAlaLeuGlnSerLeu	460
Db	1321	GGGTTACATCGTGTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA	1380
Qy	461	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys	480
Db	1381	TCATCTGAAACACTTTCGGAACCTCAGGTTATAGAAATGCCTTATGCTTACCAGTGCTGT	1440
Qy	481	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
Db	1441	GCATTTGGAGTGTGAGATGCCTATAAGATTCTAATCAATGGAATAAAGGTGACAA	1500
Qy	501	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
Db	1501	AGCAGTATGACGACCTTCATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Qy	521	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540
Db	1561	GACCTTGAGATTTCCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTACAGTCGAG	1620
Qy	541	CysSerProSerProGlyProPheLysProCysGluHisIleLeuLeuAspGlyTrpLeu	560
Db	1621	TGTTCACTTCCCAGCCGCTTCAAAACCCGTGGAACACCTGCTGTGATGCTGGCTGATC	1680
Qy	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
Db	1681	AGAAATGGAGTGTGACCATAGCAGTCTGGCACCTACTTGTAATGCTTTGGTGACTTCA	1740
Qy	581	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600
Db	1741	ACAGTTTTCAGATCCCCTCTGTACATTTCCCCCAATTAACGTTAATGGGGTCAATCGCA	1800
Qy	601	AlaValAsnMetIleuThrGlyValSerSerAlaValLeuAlaGlyValaspAlaPheThr	620
Db	1801	GCAGTGAACATGCTCAGGGAGTCTCAGTGGCCGTGCTGCTGCTGGAGTGGCTTCACT	1860
Qy	621	PheGlySerPheAlaArgHisGlyValATrPTrpGluAsnGlyValGlyCysHisValIle	640
Db	1861	TTTGGCAGCTTGCACGACATGGTGCCTCGTGGGAGAATGGGTGGTGCATGTCAAT	1920
Qy	641	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu	660
Db	1921	GGTTTTTGTGTCATTTTTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG	1980
Qy	661	GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaPropheSerSer	680
Db	1981	GAGCGTGGGTCTCTGTGAAATATCTGCAAATTTGAAACGAAGCTTCCATTTTCTAGC	2040
Qy	681	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700
Db	2041	CTGAAAGTAATCATTTTGCTCTGTGCCCTGTGGCCCTTGACCATTGGCGCAGATTCCCTG	2100
Qy	701	LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro	720
Db	2101	CTGGGTGGCAGCAAGTAGTGGCGCTCCCTCTCGCTTCTGCTTTTGGGAGGCC	2160
Qy	721	SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740
Db	2161	AGCACCATGGCTACATGGTCGCTCTCATCTTGCTCAATTCCTTTGCTTCTCATGATG	2220
Qy	741	ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp	760
Db	2221	ACCAATGCCCTACACCAAGCTCTACTGCAATTTGCAACGAGGACCTGGAGAAATATTGG	2280
Qy	761	AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
Db	2281	GACTGTCTATGGTAAACAACATTTGCCCTGTGTCTCTCACCACTGCATCTCTAAATGTC	2340
Qy	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
Db	2341	CCTGTGGCTTCTGTGCTCTCTCCTTTTAATAACCTTACATTTATCAGTCTCGAGATA	2400

Qy	801	IleLysPheIleLeuLeuValValProLeuProAlaCysIeuAenProLeuLeuTyr	820
Db	2401	ATTAAAGTTTATCCTTCCTGGGTAGTCCCACTTCCTGCATGTCTCAATCCCTTCTCTAC	2460
Qy	821	IleLeuPheAenProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal	840
Db	2461	ATCTTGTTTCAATCCTCACTTTAAAGGAGGATCTGGTGAGCCTCGAAGAACGAAACCTACGTC	2520
Qy	841	TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAspAapValGluLysGln	860
Db	2521	TGCACAGATCAAAACACCCCAAGCTTGATGTCAATTAACTCTGATGATGCGAAACACAG	2580
Qy	861	SetCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCATTTATGACCTG	2640
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2641	CCTCCCAAGTTCCTGGGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2700
Qy	901	ValAlaPheValProCysLeu	907
Db	2701	GTGGCATTTTGTGCCATGCTC	2721

RESULT 4

ADH14255	ADH14255 standard; cDNA; 2724 BP.
XX	
XX	ADH14255;
XX	
XX	11-MAR-2004 (first entry)
DT	
XX	Human HG38 cDNA.
DE	
XX	human; non-endogenous; G protein-coupled receptor; GPCR; ss; gene.
KW	
XX	Homo sapiens.
OS	
XX	US2003105292-A1.
PN	
XX	05-JUN-2003.
PD	
XX	20-SEP-2002; 2002US-00251385.
XX	
PF	
XX	26-JUN-1998; 98US-0090783P.
PR	
PR	07-AUG-1998; 98US-0095677P.
PR	13-OCT-1998; 98US-00170496.
XX	
XX	(LIAW/) LIAW C W.
PA	(BEHA/) BEHAN D P.
PA	(CHAL/) CHALMERS D T.
XX	
PI	Liaw CW, Behan DP, Chalmers DT;
XX	
XX	WPI; 2003-801247/75.
DR	
DR	P-PSDB; ADH14256.
XX	
XX	New constitutively active, non-endogenous version of an endogenous human
PT	G protein-coupled receptor for the identification of therapeutic
PT	compounds, such as agonists.
XX	
XX	Example 1; SEQ ID NO 263; 227pp; English.
PS	
XX	
CC	The invention relates to a constitutively active, non-endogenous version
CC	of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC	used for screening therapeutic compounds as inverse agonists, agonists or
CC	partial agonists. The GPCR can be also used to elucidate and
CC	understand the roles of GPCRs in normal and diseased humans. The GPCR
CC	need not be purified and isolated to be used to screen for therapeutic
CC	compounds. The utility of the GPCR as a research tool is enhanced because
CC	the role of a particular receptor can be understood before the endogenous
CC	ligand is identified. The present sequence is used in the exemplification
CC	of the present invention.
CC	

QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 901 TTTTCAACATTTACCTGAACTAAGACACACAGACTCTGAATGGTGCCTCACAAATAACTGAA 960
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 961 TTTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCCCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1081 TACAACCTTATTAGAAGATTACCCAGTTTTTCAGTCTGCCAAGAAGCTTCAGAAATTTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTTCCAGCAGTTGCTTAGCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTATTCACCCCAATGCAATTTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCTCCTAATAAAGCTGACCTATGCTCCAACTTCTGCTCTTTTCTTATTAAT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCACCTTAAAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTATGCTTTACCAAGTGGTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGAGAAATGCCCTATAGATTTCTAATCAATGAATAAAGGTGACAAAC 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTAGGACGACCTTCATAGAAGAGTGTGGAATGTTTCAAGCTCAAGATGAACGT 1560
QY 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaIleHisSerValGln 540
Db 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1621 TGTTTACCTTCCCAGGCCCTTCAAAACCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTCAGATGCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGGTCAATCGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGGTGCTGGCTGGTGGATGGATGGTTCAT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTTGGCAGCTTTTGACGACATGGTGGCTGGTGGAGAAATGGGGTGGTGGCCATGTCAAT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 1921 GGTTTTTTGTCCATTTTTCCTTCAAGAAATCATCTGTTTTCTGCTTACTCTGGAGGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

Db 1981 GAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAATCATTTTGTCTGTGCTCCCTGTGGCTTTGACCATGGCGCAGTTCCCTCG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAGATGATGGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGTACATGGTGTCTCTCATCTTGTCAATTCCTTGTCTTCTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuAsnIleTyr 760
Db 2221 ACCATTGGCTTACCAACAGCTCTACTGCAATTTGGCAAGGGGAGACCTGGAGAATATTTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2281 GACTGCTTATGGTAAACACACATTGCCCTGTGCTCTTCCACCACTGCATCTTAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuLeuThrPheIleSerProGluVal 800
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2401 ATTAAGTTTATCTTCTGTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2461 ATCTTGTTCATCTCTCACTTTAAGAGGATCTGGTGAGCCTGAGAAAGCAAACTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2521 TGGCAACAGATCAAAACACCCAAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerSerIleThrTyrAspLeu 880
Db 2581 TCTCTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2641 CTTCCAGTTCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
Db 2701 GTGGCAATTTGCCATGTCTC 2721
RESULT 5
ADI32985
ID ADI32985 standard; DNA; 2724 BP.
XX AC ADI32985;
XX DT 22-APR-2004 (first entry)
XX DE Human G protein-coupled receptor (GPCR) 49 DNA.
XX KW G protein-coupled receptor 49; GPCR; neuroprotective; neural;
XX KN endocrine system disorder; gene therapy; antisense therapy; human; ds.
XX OS Homo sapiens.
XX PN US2003235910-A1.
XX PD 25-DEC-2003.
XX PF 17-JUN-2002; 2002US-00174456.
XX PR 17-JUN-2002; 2002US-00174456.

561 ArgIleGlyValThrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db |||||
1681 AGAATTGGAGTGTGGACCATAGCAGTTCGGCACCTACTTGTAAATGCTTGTGGTCACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db |||||
1741 ACAGATTTTCAGATCCCTCTGTACATTTCCGCCATTAACCTGTAATTTGGGGTCATCGCA 1800
QY 601 AlaValAsnMetLeuThrClyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db |||||
1801 CAGATGAACATCTCACGGAGTCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrrPGLuAenGlyValGlyCysHisValIle 640
Db |||||
1861 TTGGCAGCTTTGACAGCATGGTGGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db |||||
1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
QY 661 GluArgGlyPheSerValIlySerAlaLysPheGluThrIlySerAlaProPheSerSer 680
Db |||||
1981 GAGCGTGGTCTCTGTGAAATTAATTCGAAAATTTGAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db |||||
2041 CTGAAGTAATCATTTTGTCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
Db |||||
2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db |||||
2161 AGCACATGGGGTACATGGTGGCTCTCTCAATTTGCTTCAATTTGCTTCAATTTGCTTCA 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrr 760
Db |||||
2221 ACATTTGCTTACCAACAGCTCTCTGCAATTTGGACAGGGAGACTGGAGATATTTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
Db |||||
2281 GACTGCTCTATGTTAAACACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
QY 781 ProValAlaPheLeuSerPheSerSerIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
Db |||||
2341 CCTGTGGCTTTCTGTGCTTCT 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db |||||
2401 ATTAAGTTTATCTTCTGTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db |||||
2461 ATCTGTGTTCAATCTTCACTTTAAGGAGGATCTGGTGGCTGAGAAAGCAACCTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
Db |||||
2521 TGGCAAGATCAAAACACCAAGCTTGATGTCAATTAATCTCTGATGATGTGCAAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db |||||
2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProLafyProValThrGluSerCysHisLeuSerSer 900
Db |||||
2641 CCTCCAGTTCGGTGGCATCACCAGCTTATCCAGTGACTGAGCTGCCATCTTCTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
Db |||||
2701 GTGGCATTTGTCCCATGCTCTC 2721

RESULT 6

ADO29923
ID ADO29923 standard; cDNA; 2724 BP.
XX
AC: ADO29923;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human GPCR GPR49 polynucleotide, SEQ ID NO:1025.
XX
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
transgenic mouse; neurological disorder; adrenal gland disorder;
colon disorder; intestinal disorder; cardiovascular disorder;
muscular disorder; blood disorder; immune disorder; bone disorder;
joint disorder; metabolic disorder; nutritive disorder; cancer;
kidney disorder; liver disorder; lung disorder; breast disorder;
ovary disorder; uterus disorder; prostate disorder; testis disorder;
skin disorder; stomach disorder; pancreas disorder; spleen disorder;
thymus disorder; thyroid disorder; antiparkinsonian; antidiabetic;
cytostatic; antiinflammatory; vasotropic; antidiarrhoeic; antidiabetic;
CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
virucide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
dermatological; antitumor; antithyroid; antiallergic; anorectic;
immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
gene; ss.
XX
KW Homo sapiens.
OS
XX
PN WO2004040000-A2.
XX
PD 13-MAY-2004.
XX
PF 09-SEP-2003; 2003WO-US028226.
XX
PR 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX
PA (PRIM-) PRIMAL INC.
XX
PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX
DR WPI; 2004-390329/36.
XX
PT P-PSDB; ADO29408.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
compounds that modulates diagnosing and treating disease condition
associated with GPCR dysfunction e.g. autoimmune diseases, angina
pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 1025; 542pp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
(GPCRs) and nucleic acids encoding them. The invention also relates to
sequences at least 90% identical to the GPCR proteins and nucleic acids
of the invention; methods of treating, preventing or diagnosing diseases
associated with GPCRs of the invention; methods of screening for
compounds useful in the treatment of GPCR-related diseases; a transgenic
mouse comprising a GPCR gene of the invention; a mouse comprising a
mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
from the transgenic mice; kits comprising several mice, each of which has
a mutation in a different GPCR gene of the invention; and kits comprising
probes which hybridise to GPCR polynucleotides of the invention. The
invention further discloses variants of the GPCR polypeptides and vectors
comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
be used in the diagnosis, treatment or prevention of a wide variety of
diseases including neurological disorders (e.g., Alzheimer's disease,
depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
disorders of the adrenal gland; disorders of the colon or intestine
(e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
myocardial infarction); muscular disorders; blood disorders (e.g.,
anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid

2612

Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTtpGluAanGlyValGlyCysHisValIle 640
Db |||||
Qy 1861 TTGGCAGCTTTGCACGACATGGTCCCTGGTGGGAGAAATGGGTGGTGGCATGTCATT 1920
Db |||||
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThreValAlaLeu 660
Db |||||
Qy 1921 GGTATTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGGCAGCCCTG 1980
Db |||||
Qy 661 GluArgGlyPheSerValIleValSerAlaValPheGluThrLysAlaProPheSerSer 680
Db |||||
Qy 1981 GAGCGTGGGTTCCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db |||||
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db |||||
Qy 2041 CTGAAAGTAATCATTTTGTCTCTGTGCCCTGTGGCTTGACCATGGCCGAGTTCCTCTG 2100
Db |||||
Qy 701 LeuGlyGlySerLysIleGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db |||||
Qy 2101 CTGGGTGGCAGCAGATGATGGCCCTCCCTCTCTGCTGCTTGTGCTTTTGGGAGGCC 2160
Db |||||
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db |||||
Qy 2161 AGCACCATGGGTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db |||||
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db |||||
Qy 2221 ACCATGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATTTTGG 2280
Db |||||
Qy 761 AspCysSerMetValIleHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
Db |||||
Qy 2281 GACTGCTCTATGTAACACACATTTGCCCTTTGCTCTTCAACCACTGCATCTCAAACTGC 2340
Db |||||
Qy 781 ProValAlaPheLeuSerPheSerSerIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
Db |||||
Qy 2341 CCGTGGCTTTCTGTGCT 2400
Db |||||
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db |||||
Qy 2401 ATTAAGTTTATCTTCTGTGGTGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Db |||||
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db |||||
Qy 2461 ATCTGTTTCAATCTCTCACTTTAAGGAGGATCTGCTGAGCCTGAGAAAGCAACCTACGTC 2520
Db |||||
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db |||||
Qy 2521 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTGTGATGTGCAAAACAG 2580
Db |||||
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db |||||
Qy 2581 TCTGTGACTCACTCAAGCTTTGTAACCTTTTACCTTACAGCTCCAGCATCATCTATGACCTG 2640
Db |||||
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db |||||
Qy 2641 CTCTCCAGTTCTGGTCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
Db |||||
Qy 901 ValAlaPheValProCysLeu 907
Db |||||
Qy 2701 GTGGCATTTGTCCTATGCTC 2721
Db |||||

RESULT 7
ADN39796
ID ADN39796 standard; cDNA; 2880 BP.
XX
AC ADN39796;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C168.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;

KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-0355250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-036809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-037246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-039775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX RI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX WPI; 2003-468649/44.
XX DR P-PSDB; ADN40013.
XX PT Determining the presence or absence of a pathological cell in a patient,
XX PT useful for diagnosing, prognosing or treating cancer, comprises detecting
XX PT a nucleic acid in a biological sample.
XX PS Claim 8; SEQ ID NO C168; 1385pp; English.
XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX CC whose expression is upregulated or downregulated in specific cancers or
XX CC other diseases such as angiogenic or fibrotic disorders, and to methods
XX CC of determining the presence or absence of a pathological cell in a
XX CC patient by detecting a nucleic acid at least 80% identical to those of
XX CC the invention or by detecting a polypeptide of the invention. The
XX CC invention also relates to expression vectors and host cells comprising a
XX CC nucleic acid of the invention; antibodies which specifically bind a
XX CC polypeptide of the invention; use of such antibodies for drug targeting;
XX CC and methods of screening for modulators of activity or expression of the
XX CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX CC antibodies and methods are useful for diagnosing, prognosing and treating
XX CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX CC neovascularisation syndromes, scarring and uterine fibroids. They may
XX CC also be useful in wound healing and in contraception. The present
XX CC sequence represents a nucleic acid sequence of the invention.
XX SQ Sequence 2880 BP; 710 A; 754 C; 594 G; 822 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
DB 649 ATACCAGACATGATCCCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATACAAAT 708
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 709 AGAATCCACTCCCTGGGAAGAAGATGCTTTGATGGGCTCCACAGCTTAGAGACTTAGAT 768
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 769 TTAATAATACAATAACCTTGATGAATCCCACTGCAATTAGGACACTCTCCAACCTTAAA 828
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 829 GAACTAGGATTTTCATAGCAACAATATCAGTGCATACCTGAGAAAGCANTTTGAGCAAC 888
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 889 CCTTCTCTTATACATAATCTTCTATGACATCCCATCCATTTGTTGGGAGATCTGCT 948
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 949 TTTCAACATTTTAACTGAACCTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 1008
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 1009 TTTCTCGATTTTAACTGAACTGCAACCTGGAGAGTCTGACTTTTAACTGGAGCACAGATC 1068
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1069 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAGTGTCTAGATCTGCT 1128
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1129 TACAACCTTATAGAGATTTACCAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1188
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1189 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCAGCAGTGTGCTTAGACCTC 1248
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1249 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTTATTCACCCCAATGCAATTTTCCACT 1308
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
DB 1309 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACCTCTCTGCTGCTTTTCTATAACT 1368
QY 441 GlyLeuHisGlyLeuThrHisLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1369 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1428
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1429 TCATCTGAAACCTTTCCAGACTCAAGGTATAGAAATGCCATTATGCTTACCAGTGCTGT 1488
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1489 GCATTTGGAGTGTGAGAAATGCCATATAAGATTTCTAATCANTGAATAAGGTGACAC 1548
QY 501 SerSerMetAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1549 AGCAGTATGAGGACCTTCATAGAAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
QY 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1609 GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTTACGTGAC 1668
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1669 TGTTCACTTCCCGAGGCCCTTCAACCCCTGTGAACACTGCTGTGAGTGGCTGGCTGATC 1728
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580

RESULT 9
ADR67869

DB 1729 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATCCTTTGGTGACTTCA 1788
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1789 ACAATTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTAAATGGGTGTCATCGCA 1848
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1849 GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGGTTCCT 1908
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1909 TTTGGCAGCTTTTGCACGACATGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATT 1968
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1969 GGTTTTGTTCATTTTGTCTTCAGATCATCTGTTCCTGCTTACTCTGGCAGCCCTG 2028
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 2029 GAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2088
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2089 CTGAAGATTAATCATTTTGTCTCTGTGCCCTGTGGCCTTGACCATGGCCCGAGTTCCTCTG 2148
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
DB 2149 CTGGGTGGCAGCAGATAGGCGCCTCCCTCTCTGCTGCTGCTTTGGCTTTGGGAGCC 2208
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
DB 2209 AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2268
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGlnIleTyr 760
DB 2269 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATATTGG 2328
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2329 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCCACCACTGCATCTCTAACTGC 2388
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2389 CCTGTGGCTTTCTGTCT 2448
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2449 ATTAAGTTTATCTCTGTGGTAGTCCCACTTCTGTCATGTCTCAATCCCTTCTCTAC 2508
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2509 ATCTGTTCAATCTCTCACTTTTAAGAGGATCTGGTAGGCTGAGAAAGCAAACTACGTC 2568
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
DB 2569 TGGCAAGATCAAAACACCCAGCTTGATGTCAAATAACTCTGTGATGATGCGAAAAACAG 2628
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2629 TCTGTGATCTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCTCTATGACCTG 2688
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2689 CTTCCAGTTCCTGCTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCT 2748
QY 901 ValAlaPheValProCysLeu 907
DB 2749 GTGGCATTTGTCCCATGCTC 2769

Db 1158 CTAAGACATAATGAATCTACGAATTAAGTTGACATTTTCAGCAGTTGCTTACGCTC 1217
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1218 CGATCGCTGAATTTGGCTGGACAAATAATGCTATTATTACACCCCAATGCAATTTCCACT 1277
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1278 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACCTCGCTGCTTTTCCCTATAACT 1337
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1338 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCCCTTACAGAGCTTGATA 1397
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1398 TCATCTGAAACTTTCCAGACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTCT 1457
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1458 GCATTTGGAGTGTGTGAGAATGCTATAAGATTCTTAATCAATGGAATAAAGGTGACAAC 1517
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1518 AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1577
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1578 GACCTTGAAAGATTCTGCTGTGACTTTGAGGAAGACCTGAAGCCCTTCATTGAGTGCAG 1637
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1638 TGTTTACCTTTCCACAGCCCTTCAAACCTGTGAACACCTGCTTCAATGGCTGATC 1697
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1698 AGAATTTGGAGTGTGGACCATAGCAGTCTGGCACCTTACTTGTAAATGGGTGACTTCA 1757
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1758 ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAATGTTAAATGGGTGATCGCA 1817
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1818 GCAGTGAAACATGCTCACGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGGTTCAT 1877
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1878 TTTGGCAGCTTTGACAGCATGCTGCTGTGGAGAAATGGGTGCTTGGCCATGTCATT 1937
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 1938 GGTTTTGTGTCATTTTTCCTTTCAGAAATCATCTGTTTTCCTGCTTACTTGGCAGCCCTG 1997
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1998 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAACGNAAGTCCATTTTCTACG 2057
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
Db 2058 CTGAAAGTAAATCATTTTGTGTCCTGCTGGCCTTGCACCATGGCCGAGTTCCTCCCTG 2117
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2118 CTGGGTGGCAGCAGATGCGGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGCCC 2177
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2178 AGCACCATGGGTACATGCTGCTCATCTTCTGCTCAATTCCTTTGCTTCTCTCATGAG 2237
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 2238 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGCCTGGAGATATTTGG 2297
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2298 GACTGCTCTATGTAACACACATTTGCCCTGTTGCTCTTCCACCACTGCATCTTAAACTGC 2357
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2358 CCGTGGCTTCTTGTGCTTCTCTCTTTAATAAACCTTACATTTTATCATGCTCCTGAAGTA 2417
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2418 ATTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2477
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2478 ATCTTGTTCATCTCACCTTTAAGGAGGATCTGGTGAGCTTGAGAAAGCAACCTACGTC 2537
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2538 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTTAACTCTGATGATGTGCAAAACAG 2597
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2598 TCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCATCTATGACCTG 2657
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2658 CTCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2717
Qy 901 ValAlaPheValProCysLeu 907
Db 2718 GTGGCATTTGTCCCATGTCTC 2738
RESULT 10
ID ADN39627 standard; cdna; 3032 BP.
XX
AC ADN39627;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:A227.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.

QY 481 AlaPheGlyValCysGluAsnAlaThrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1641 GCATTTGGAGTGTGAGAAATGCTATAAGATTCTTAATCAATGGAATAAGGTGACAAC 1700
QY 501 SerSerMetAspAspLeuHisLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1701 AGCAGTATGACGACCTTCATAGAAGAAGATGCTGGAAATGTTTCAGGCTCAGATGAACGT 1760
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1761 GACCTTGAAGATTCTGCTGTGACTTGTAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG 1820
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1821 TGTTCACCTTCCCGAGCCCTTCAAAACCTGTGAAACACTGCTTGTATGGCTGGCTGATC 1880
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1881 AGNATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATTCA 1940
QY 581 ThrValPheArgSerProLeuTrpIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1941 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGTTAATGGGGTCATCGCA 2000
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2001 CGAGTGAACATGCTACGGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGT 2060
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 2061 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 2120
QY 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660
Db 2121 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2180
QY 661 GluArgGlyPheSerValLysTrpSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2181 GAGCGTGGGTCTCTGTGAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2240
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
Db 2241 CTGAAGATTAATCATTTTGTCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300
QY 701 LeuGlyGlySerLysTrpGlyAlaSerProLeuCysLeuLeuProPheGlyGluPro 720
Db 2301 CTGGGTGGCAGCAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTGGGGAGCCC 2360
QY 721 SerThrMetGlyTrpMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2361 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2420
QY 741 ThrIleAlaTrpThrLysLeuTrpCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2421 ACCATGGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGG 2480
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2481 GACTGCTCTATGTAACACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2540
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2541 CCTGTGGCTTCTTGTGCTTCT 2600
QY 801 IleLysPheIleLeuValValProLeuProAlaCysLeuAsnProLeuLeuTrp 820
Db 2601 ATTAAAGTTTATCTCTCTGTTGGTAGTCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2660
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTrpVal 840
Db 2661 ATCTGTTCAATCTCACTTTAAGGAGGATCTGGTGAGCTCGAGAAAGCAACCTACGTC 2720
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860

Db 2721 TGACCAAGATCAAAACACCAAGCTTGTCAATTAATCTGATGATGTCGAAACAG 2780
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTrpAspLeu 880
Db 2781 TCTGTGACTCAACTCAAGCTTGGTAACTTTACCGCTCCAGCATCACTTATGACCTG 2840
QY 881 ProProSerSerValProSerProAlaTrpProValThrGluSerCysHisLeuSerSer 900
Db 2841 CCTCCAGTTCCGTGCCATCAGCAGCTTATCCAGTGACGTGAGAGCTGCCATCTTCTCT 2900
QY 901 ValAlaPheValProCysLeu 907
Db 2901 GTGGCATTTGTCCCATGTCTC 2921
RESULT 11
ADN39530
ID ADN39530 standard; cDNA; 3032 BP.
XX
AC ADN39530;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: A130.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; poriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372248P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi P;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39531.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO A130; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a nucleic acid sequence of the invention.

```

Alignment Scores:
Pred. No.:          0
Score:              907.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:        100.00%
DB:                  11
Length:              3032
Matches:             907
Conservative:        0
Mismatch:             0
Indels:              0
Gaps:                0

US-10-751-736-84 (1-907) x ADN39530 (1-3032)

```


Db 549 GCTCTGACATATCTCCAGGAGCATTCACCTGGCCCTTACAGCTCTTAAAGTCTTATG 608
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 609 CTGCAGATAATCAGCTAAGACACGTAACCCACAGAAAGCTCTGCAGATTTTGGGAAGCCTT 668
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 669 CAATCCCTGCGTCTGATCTAACCACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 728
Qy 161 LeuHisSerLeuArgHisIleuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Db 729 CTGCATCTCCCTGAGGACCTGTGGCTGGATGCAATGCGTTAAACAGAAATCCCGCTCCAG 788
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Db 789 GCTTTTAGAAGTTTATCGGCAITTCGAAGCCATGACCTTGGCCCTGAACAAATAACACCAC 848
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisIleuHisAsnAsn 220
Db 849 ATACCAAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATCAACAAT 908
Qy 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 909 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTAGAGACTTTAGAT 968
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 969 TTTAAATTACAATAAACCCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA 1028
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 1029 GAACTAGGATTTTCATAGCAACAATATCAGTGCATACCTCGAGAAAGCAATTTGTAGGCAAC 1088
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1089 CCTTCTCTTATTACAAATACATTTCTATGAAATCCATCCATCCATTTGTTGGAGACTGCT 1148
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1149 TTTCAACATTTACCTGAACATAAGAACACACTGACTCTGAATGGTGCCTCACAATAACTGAA 1208
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1209 TTTCTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1268
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1269 TCATCTCTTCTCCAAACCGTCTGCAATCAGTATGACTTAATCTCCAGTGTCTAGATCTGCT 1328
Qy 361 TyrAsnLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1329 TACAACCTATTAGAAGATTTACCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1388
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1389 CTAAGACATAAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGACCTC 1448
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1449 CGATCCCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTATCAACCCAAATGCTTTTCCATAACT 1508
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1509 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTGTCCGCTTTTCTCTATAACT 1568
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1569 GGGTTACATGGTTTAACTCATTAAATTAACAGGAAATCATGCGCTTACAGAGCTTGATA 1628
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480

Db 1629 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTTACCAGTCTGT 1688
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1689 GCATTTGGAGTGTGTGAGAATGCCCTATAAGATTTCTTAATCAATGAATAAAGGTGACAAC 1748
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1749 AGCATATGGACAGCTTCATAGAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT 1808
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1809 GACCTTGAAGATTTCTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGTCAG 1868
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1869 TGTTCACTCTCCCAAGGCCCTTCAAAACCTGTGTGAACACCTGCTTGATGGCTGGCTGATC 1928
Qy 561 ArgIleGlyValIleThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1929 AGAATGGAGTGTGGACCATAGCAGTTCTGGACATTTCTGTAATGCTTTGGTGACTTCA 1988
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuLeuIleGlyValIleAla 600
Db 1989 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGTCATCGCA 2048
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2049 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCTGGCTGGCTGGTGGATCGGTTCACT 2108
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 2109 TTTGGACGCTTTGACAGCATGCTGCTGGTGGAGAAATGGGTTGGTGGCTGATGCTCAT 2168
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 2169 GGTTTTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2228
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2229 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCCTAGC 2288
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2289 CTGAAGATTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2348
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2349 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTTGTGCTTGTGCTGCTGCTGCT 2408
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2409 AGCACCATGGGTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2468
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2469 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGAGCTGGAGATATTGG 2528
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2529 GACTGCTCTATGGTAAACACATTTGCTGCTTGTGCTTCTTCAACAACTGCATCTCTAACTGC 2588
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2589 CCTGTGGCTTCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2648
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2649 ATTAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2709 ATCTTGTTCATCTCCTCCTTTTAAAGGAGGATCTGGTGGCTTGAGAAAGCAACCTACGTC 2768

Qy	841	TrpThrArgSerIysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln	860
Db	2769	TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTAACTCTGATGATGCGAAAAACAG	2828
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2829	TCCTGTGACTCAACTCAAGCCCTGGTAACCTTTACCAGCTCCAGCATCATTTATGACCTG	2888
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2889	CCTCCAGTTCCTGGCCATCACAGCTTATCCAGCTGACTGAGCTGCCATCTTTCCTCT	2948
Qy	901	ValAlaPheValProCysLeu	907
Db	2949	GTGGCATTTGTCCCATGTCTC	2969

RESULT 13

ADL12472
ID ADL12472 standard; cDNA; 4570 BP.
XX
XX ADL12472;
XX
XX ADL12472;
XX
XX 06-MAY-2004 (first entry)
XX
XX Human steroid-induced C3A liver cell cDNA #201.
DE
XX
XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
KW
XX Homo sapiens.
OS
XX US6673549-B1.
PN
XX 06-JAN-2004.
PD
XX
XX 12-OCT-2001; 2001US-00976594.
PF
XX
XX 12-OCT-2000; 2000US-0240409P.
PR
XX
XX (INCY-) INCYTE CORP.
PA
XX
XX Furness LM, Buchbinder JL;
PI
XX
XX WPI; 2004-068610/07.
DR
XX
XX
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
XX Claim 1; SEQ ID NO 201; 141pp; English.
PS
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and a
CC least one molecule or compound. The molecules or compounds are regulator
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX

|||||
1162 TTTCAACATTTACCTGAACCTAGAACACCTGACTCTGAATGGTGGCTCCCAATAACTGAA 1221
QY PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
|||||
1222 TTTCTGATTTTAACTGGAACCTGAAACCTGGAGAGCTGACTTTAACTGGAGCACAGATC 1281
Db
QY SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db TCATCTCTCTCTCAACACCGTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCT 1341
QY TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db TACAACCTATTAGAGATTACCCAGTTTTCAGTCTGCCAAGAGCTTCAGAAAATTGAC 1401
QY LeuArgHisAsnGluIlePyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1461
QY ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTATTCACCCCAATGCAATTTCCACT 1521
QY LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db TTGCCATCCCTAATAAAGCTGACCTATCGTCAACCTCCTGCTCTTTTCCATATACT 1581
QY GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db GGGTTACATGGTTTAACTCACCTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1641
QY SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATATGCTTTACCAGTGGCTGT 1701
QY AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db GCATTTGGAGTGTGAGAAATGCCATATAAGATTTCTAATCAATGGAATAAAGTGCACAC 1761
QY SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1821
QY AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGGCAG 1881
QY CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db TGTTTACCTTCCCAGGCCCTTCAAACCTGTGGAACACCTGCTGTGATGGCTGGCTGATC 1941
QY ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA 2001
QY ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTTAAACTGTTAATTTGGGTGCTATCGCA 2061
QY AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db GCAGTGAACATGCTCAGCGAGTCTCCAGTGGCGTGTGGCTGGTGTGATCGGTTCAT 2121
QY PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db TTTGGCAGCTTTGCACGACATGGTGGCTGGTGGAGATGGGGTTGGTTGCCATGTCAAT 2181
QY GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrIleAlaLeu 660
Db GGTTTTGTGCTTTCATTTTCTCAGAAATCATCTGTTTCTCCTGCTTACTCTGGAGCCCTG 2241
QY GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

Db 2242 GAGCGTGGTTCCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2301
QY LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db CTGAAGTAATCATTTTGTCTGTGCCCTGTGGCTTTGACCATGGCGCAGTTCCCTCG 2361
QY LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db CTGGGTGGCAGAGTATGGGCCCTCCCTCTCTGCGCTGCCCTTGGCTTTGGGGAGGCC 2421
QY SerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db AGCACATGGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTTGCTTCTCATGATG 2481
QY ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db ACCATTTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATATTGG 2541
QY AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db GACTGCTCTATGGTAAACACATTTGCCCTGTGCTTTCACCAACTGCATCCTAACTGC 2601
QY ProValAlaPheLeuSerPheSerSerLeuLeuLeuAsnLeuThrPheIleSerProGluVal 800
Db CCTGTGGCTTTCTGTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2661
QY IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db ATTAAGTTTATCTCTTCTGTGGTAGTCCCACCTTCCCTGATGCTCAATCCCCCTTCTCTAC 2721
QY IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db ATCTTGTTCATCTCACTTTAAGAGGATCTGGTGAGCCTGAGAAAGCAAACTACCTACGTC 2781
QY TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTAATCTCTGATGATGTCGAAAAACAG 2841
QY SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db TCCTGTGACTCAAACTCAAGCCTTGATAACCTTTTACCAGCTCCAGCATCACATTATGACCTG 2901
QY ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db CCTCCAGTTCGGTGCCATCACACGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2961
QY ValAlaPheValProCysLeu 907
Db GTGGCATTTGTCCCATGTCTC 2982
RESULT 14
AA23981
ID AAX23981 standard; DNA; 3297 BP.
XX
AC AAX23981;
XX 25-JUN-1999 (first entry)
XX
DE Human HG38 DNA.
XX
KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
KW endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator; ss.
OS Homo sapiens.
XX
PN W09915660-A1.
XX
PD 01-APR-1999.
XX
PF 24-SEP-1998; 98WO-US019979.
XX
PR 24-SEP-1997; 97US-0059863P.


```
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db |||||
2221 ACCATGGCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATATTTGG 2280
QY 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780
Db |||||
2281 GACTGCTCTATGAA-AAAACACATTGCCCTGTTGCTCTTCACCAACTGCATCCTAAACTG 2339
QY 780 sProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
Db |||||
2340 CCCTGTGGCTTTCCTGCTCTCTCTCTTAATAAACCTTACATTTATCAGTCTGAAAGT 2399
QY 800 ILeLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
Db |||||
2400 AATTAAGTTTATCCTTCTGGTGGTAGTCCCACCTTCTGCGATGCTCAATFCCCCCTTCTCTA 2459
QY 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyxVa 840
Db |||||
2460 CATCTTGTTCATCTCCTCATTAAAGAGGATCTGGTGAGCCTGAGAAAGCAACCTACGT 2519
QY 840 ITrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGI 860
Db |||||
2520 CTGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGATGATGTCGAAAAACA 2579
QY 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyxAspLe 880
Db |||||
2580 GTCCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACCAAGCTCCAGCATCAGTTATGACCT 2639
QY 880 uProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSe 900
Db |||||
2640 GCCTCCAGTTCGGTGCCATCACCAAGCTTATCCAGTACTGAGAGCTGCCATCTTTCCTC 2699
QY 900 rValAlaPheValProCysLeu 907
Db |||||
2700 TGTGGCATTGTGCCATGTCTC 2721
```

Search completed: July 12, 2005, 08:18:07
Job time : 1325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: July 12, 2005, 07:56:07 ; Search time 2965 Seconds
(without alignments)
1920.553 Million cell updates/sec
Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLKGLLSLVLLQLAT.....PAYPTVTSCHLSLVAFVCL 907

Scoring table: OLIGO 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6330945 seqs, 3139162390 residues

Word size: 1

Total number of hits satisfying chosen parameters: 12650801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+p2n.model
-O=/cgn2_1/USPTO.spool/US10751736/runat_12072005_083643_22228/app_query.fasta_1.1095
-DB=Published Applications NA -QWTF=fastap -SUFFIX=oli.Tnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MAIRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USRP=US10751736 @CGN 1.1 859 @runat_12072005_083643_22228
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCX=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
- 22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	907	100.0	2724	15	US-10-251-385-263
2	907	100.0	2724	15	US-10-225-567A-421
3	907	100.0	2724	17	US-10-174-456-4
4	907	100.0	2724	20	US-10-751-736-21
5	907	100.0	2880	17	US-10-285-027-1114
6	907	100.0	2880	21	US-10-482-029-157
7	907	100.0	2880	21	US-10-651-237-48
8	907	100.0	2880	21	US-10-782-413-48
9	907	100.0	3032	17	US-10-295-027-848
10	907	100.0	3032	17	US-10-295-027-945
11	820	90.4	2724	15	US-10-251-385-277
12	806	88.9	3438	21	US-10-505-486-205
13	621	68.5	2651	17	US-10-295-027-483
14	621	68.5	2651	17	US-10-295-027-1113
15	621	68.5	2651	17	US-10-173-999-27
16	606	66.8	2082	21	US-10-851-470-3
17	232	25.6	723	15	US-10-313-542-220
18	197	21.7	1790	17	US-10-264-049-957
19	23	2.5	687	13	US-10-027-632-2009
20	23	2.5	687	13	US-10-027-632-2009
21	22	2.4	576	13	US-10-027-632-74159
22	22	2.4	576	13	US-10-027-632-299355
23	22	2.4	576	17	US-10-027-632-74159
24	22	2.4	576	17	US-10-027-632-299355
25	19	2.1	60	10	US-09-908-975-19787
26	18	2.0	2208	10	US-09-851-595-9
27	18	2.0	2208	18	US-10-664-667-9
28	18	2.0	2487	14	US-10-270-336-1
29	18	2.0	2711	10	US-09-851-595-7
30	18	2.0	2711	18	US-10-664-667-7
31	18	2.0	2901	10	US-09-851-595-3
32	18	2.0	2901	10	US-09-851-595-12
33	18	2.0	2901	18	US-10-664-667-3
34	18	2.0	2901	18	US-10-664-667-12
35	18	2.0	2988	18	US-10-302-172-723
36	18	2.0	3119	15	US-10-225-567A-580
37	18	2.0	3273	17	US-10-295-027-1119
38	18	2.0	3273	20	US-10-783-528-43
39	18	2.0	3325	17	US-10-331-496A-52
40	18	2.0	3349	14	US-10-270-336-4
41	18	2.0	3381	10	US-09-970-944-5
42	18	2.0	3438	14	US-10-176-847-89
43	18	2.0	3443	17	US-10-331-496A-94
44	18	2.0	3492	10	US-09-851-595-10
45	18	2.0	3492	18	US-10-664-667-10

ALIGNMENTS

RESULT 1
US-10-251-385-263
; Sequence 263, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13

```
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 263
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-263

Alignment Scores:
Pred. No.: 0          Length: 2724
Score: 907.00        Matches: 907
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 15              Gaps: 0

US-10-751-736-84 (1-907) x US-10-251-385-263 (1-2724)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuGlnLeuAlaThr 20
DB 1 ATGGACACCTCCCGGCTGGTGCTCTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 60

QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB 61 GGGGGCAGCTCTCCCGAGGCTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATTC 120

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB 121 GAGCCGAGCGGAGGATGTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180

QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB 181 CCTTCCACCTCAGGCTCTTACCTCTTACCTAGACCTCAGTATGACACACACTCAGTCAG 240

QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuLeuAlaGlyAsn 100
DB 241 CTGCTCCCGAATCCCTCGCCAGTCTCGCTTCTCTGGAGGAGTTACGCTCTTGGGGAAAC 300

QY 101 AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMet 120
DB 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGCGCTTTACAGTCTTAAAGTCTTATG 360

QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB 361 CTGCAGATATATCAGCTAAGACACGTAACCCACAGAGGCTCTGCAGAAATTTGGAGCCTT 420

QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB 421 CAATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTCAGTGGC 480

QY 161 LeuHisSerLeuArgHisLeuThrLeuAspAspAsnAlaLeuThrGluIleProValGln 180
DB 481 CTGCAATCCCTGAGGACCTGTGGTGGATGACAAATGGCTTAAACAGAAATCCCGTCCAG 540

QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
DB 541 GCTTTTAGAAGTTATCGGCATTCAGCCATGACTTGGCCCTCGACAAATAACACCCAC 600

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisHisAsnAsn 220
DB 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660

QY 221 ArgIleHisSerLeuGlyIysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 661 AGAATCCCATCTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720

QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 721 TTAATATACAAATACCTTGATGAATTTCCCACTGCAATTAGAGACTCTCCAAACCTTAAA 780

QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluIysAlaPheValGlyAsn 280
DB 781 GAACTAGGATTTTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCATTTGTAGGCAAC 840
```

```
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 841 CCTTCTCTTATTACATACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900

QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 901 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGCTGCTCACAATAACTGAA 960

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGlnSerLeuThrLeuThrGlyAlaGlnIle 340
DB 961 TTTCTGCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC 1020

QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTCTCTCAAAACGCTCTGCAATCAGTTTACCTAATCTCCAAAGTCTGATCTGCT 1080

QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAAGATTTTACCAGTCTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140

QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAAGACATAATAGAAATCTACGAAATTAAGTTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200

QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGGAATTTGGCTTGGAAACAAAATGCTATTATTATCCACCCCAATGCAATTTCCACT 1260

QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGCCATCCCTAAATAGAGCTGACCTATCGTCCAACCTCTCTGCTCTTTTCCATAACT 1320

QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTTACATGTTTAACTCACTTAAATTAACAGAGAAATCATGCTTACAGAGCTTGATA 1380

QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTCTGT 1440

QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGCAATAAGGTGACAAAC 1500

QY 501 SerSerMetAspAspLeuHisLysIleAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGACGACCTTTCATAAGAAAGATCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560

QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGAGAGAACCTGAAAGCCCTTCATTCAGTGGCAG 1620

QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleIle 560
DB 1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680

QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATGACAGTCTGACACTTACTTGTAAATGCTTTGGTGAATCTCA 1740

QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTAAATTTGGGGTCAATCGCA 1800

QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCGCTGGCTGGCTGGTGGATGGATGCTCACT 1860

QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGGCAGCTTTGACACGACATGGTGGTGGTGGGAGAAATGGGGTTGGTGGCATGCTCAT 1920
```



```
Db      ||||| 841 CCTTCTCTATTACAAATCATTTCTATGACAAATCCATCCAAATTTGTTGGAGATCTGCT 900
Qy      ||||| 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAenGlyAlaSerGlnIleThrGlu 320
Db      ||||| 901 TTTTCAACATTTTACCTGAACATAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA 960
Qy      ||||| 321 PheProAenLeuThrGlyThrAlaAenLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db      ||||| 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
Qy      ||||| 341 SerSerLeuProGlnThrValCysAenGlnLeuProAenLeuGlnValLeuAenLeuSer 360
Db      ||||| 1021 TCATCTCTCTCAAACTGCTGCAATCAGTTACCTTAATCTCCAAGTCTGATCTGTCT 1080
Qy      ||||| 361 TyrAenLeuLeuGluAenLeuProSerPheSerValCysGlnLysLeuGlnLysIleAenP 380
Db      ||||| 1081 TACAACCTATTAGAAGATTTTACCAGTCTTCCAGTCTGCCAAAGCTTTCAGAAAATTGAC 1140
Qy      ||||| 381 LeuArgHisAenGluIleTyGluIleLysValAenThrPheGlnLeuLeuSerLeu 400
Db      ||||| 1141 CTAAGACATAATGAATCTACGAAATTAAGATTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Qy      ||||| 401 ArgSerLeuAenLeuAlaTriAenLysIleAlaIleIleHisProAenAlaPheSerThr 420
Db      ||||| 1201 CGATCGCTGAATTTGGCTTGGACAAATTCGTAATTTACCCCAATGCATTTTCCACT 1260
Qy      ||||| 421 LeuProSerLeuIleLysLeuAenLeuSerSerAenLeuLeuSerSerPheProIleThr 440
Db      ||||| 1261 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAACTCTCTGCTGCTTTTCCATAACT 1320
Qy      ||||| 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAenHisAlaLeuGlnSerIleuile 460
Db      ||||| 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATTCATGCCCTTACAGAGCTTGATA 1380
Qy      ||||| 461 SerSerGluAenPheProGluLeuLysValIleGluMetProTyAlaTyGlnCysCys 480
Db      ||||| 1381 TCATCTGAATACTTTCCAGAACTCAAGGTATAGAATGCCCTTATGCTTACCAGTCTGT 1440
Qy      ||||| 481 AlaPheGlyValCysGluAenAlaTyLysIleSerAenGlnTrpAenLysGlyAenPasn 500
Db      ||||| 1441 GCATTTGAGTGTGTGAGAATGCCCTATAGAATTTCTAATCAATGAATAAAGGTGACAAAC 1500
Qy      ||||| 501 SerSerMetAenAenLeuHisLysIleValAenPheGlnAlaGlnAenGluArg 520
Db      ||||| 1501 AGCAGTATGAGACGACCTTCAAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy      ||||| 521 AspLeuGluAenPheLeuLeuAenPheGluGluAenLeuLysAlaLeuHisSerValGln 540
Db      ||||| 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Qy      ||||| 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAenGlyTriPheIle 560
Db      ||||| 1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTGTGATGGTGGCTGATC 1680
Qy      ||||| 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAenAlaLeuValThrSer 580
Db      ||||| 1681 AGAATGGAGTGTGACCATAGCAGTCTTGGCAGCTTACTTGAATGCTTTGGTGACTTCA 1740
Qy      ||||| 581 ThrValPheArgSerProLeuTyLysSerProIleLysLeuLeuIleGlyValIleAla 600
Db      ||||| 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTAATTTGGGGTCAATCGCA 1800
Qy      ||||| 601 AlaValAenMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAenAlaPheThr 620
Db      ||||| 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Qy      ||||| 621 PheGlySerPheAlaArgHisGlyAlaTriPheGluAenGlyValGlyCysHisValIle 640
Db      ||||| 1861 TTTGGCAGCTTGGACACGATGGTCCCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGT 1920
Qy      ||||| 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
```

```
Db      ||||| 1921 GGTTTTTTGTCCATTTTGTGCTTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Qy      ||||| 661 GluArgGlyPheSerValLysTySerAlaLysPheGluThrLysAlaProPheSerSer 680
Db      ||||| 1981 GAGCGTGCGTCTCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
Qy      ||||| 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db      ||||| 2041 CTGAAGATTAATCATTTTGTCTGTGCGCTGTGCGCTTGACCATGGCCGCGAGTTCCTCCCTG 2100
Qy      ||||| 701 LeuGlyGlySerLysTyGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db      ||||| 2101 CTGGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCGCTGCGCTTTGGGAGAGCC 2160
Qy      ||||| 721 SerThrMetGlyTyThrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db      ||||| 2161 AGCACCATGGCTACATGCTGCTCATCTTGTCTCAATTCCTTTGCTTGTCTCATGATG 2220
Qy      ||||| 741 ThrIleAlaTyThrLysLeuTyCysAenLeuAenLysGlyAenLeuGluAenIleTrp 760
Db      ||||| 2221 ACCATGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTTGG 2280
Qy      ||||| 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAenCysIleLeuAenCys 780
Db      ||||| 2281 GACTGCTATATGGTAAACACATTTGCCCTGTGCTTTCACCAACTGCATCTCTAACTGC 2340
Qy      ||||| 781 ProValAlaPheLeuSerPheSerSerLeuIleAenLeuThrPheIleSerProGluVal 800
Db      ||||| 2341 CCTGTGGCTTTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Qy      ||||| 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAenProLeuLeuTy 820
Db      ||||| 2401 ATTAAGTTTATCTCTTGTGTTAGTCCCATCTCTGCAATGCTCAATCTCTCTCTCTAC 2460
Qy      ||||| 821 IleLeuPheAenProHisPheLysGluAenValSerLeuAenGlyLysGlnThrTyVal 840
Db      ||||| 2461 ATCTTGTTCATCTCACTTTTAAAGGAGATCTGTGAGCTTGAGAAACCAACTACGTC 2520
Qy      ||||| 841 TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAenAenAenAenAenAen 860
Db      ||||| 2521 TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTAATCTGATGATGTCGAAAAACAG 2580
Qy      ||||| 861 SerCysAenSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyAspLeu 880
Db      ||||| 2581 TCCGTGACTCACTCAAGCCTTGTGTAACCTTTACAGCTCCAGATCACTATGACCTG 2640
Qy      ||||| 881 ProProSerSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer 900
Db      ||||| 2641 CCTCCAGTTCCGTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTCT 2700
Qy      ||||| 901 ValAlaPheValProCysLeu 907
Db      ||||| 2701 GTGGCAATTTGTCCCATGTCTC 2721
```

RESULT 3

US-10-174-456-4

; Sequence 4, Application US/10174456

; Publication No. US20030235910A1

; GENERAL INFORMATION:

; APPLICANT: Susan M. Pfeifer

; TITLE OF INVENTION: ANTISENSE MODULATION OF G PROTEIN-COUPLED RECEPTOR 49 EXPRESSION

; FILE REFERENCE: RIG-0374

; CURRENT APPLICATION NUMBER: US/10/174,456

; CURRENT FILING DATE: 2002-06-17

; NUMBER OF SEQ ID NOS: 139

; SEQ ID NO 4

; LENGTH: 2724

; TYPE: DNA

; ORGANISM: H. sapiens

; FEATURE:

; NAME/KEY: CDS

Db 901 TTTCAACATTTACCTGAACCTAAGAACACCTGACTCTGAATGGTGGCTCAAAATAACTGAA 960
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyValaGlnIle 340
Db 961 TTTCTCTGATTTAACTTGGAACTGCAAACTGGAGAGCTGACTTTAACTTGGAGCACAGATC 1020
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAAACCGCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCT 1080
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLeuLeuGlnIleAsp 380
Db 1081 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAATAGAC 1140
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAAGACATAATGAATCTACGAATTAAGATTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCACITTAATAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAATCTTTCCAGAACTCAAGGTATAGAAATGCCTTATGCTTACCAGTGTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAATGCCCTATAGAAATTTCTTAATCAATGAATAAAGGTGACAA 1500
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGGAGCGCTTCATAGAAGAGATGCTGGAAATGTTTTCAGGCTCAAGATGACGT 1560
Qy 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAGATTTCTCTGCTGACTTTTCAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Db 1621 TGTTCCACCTTCCCAGGCGCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGTGATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATGGAGTGTGACCATGACAGTCTGGCAGCTTACTTGTAAATGCTTTGTGACTTCA 1740
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGGACAGCATGCTGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTTTGTCCATTTTGTCTCAGAACTCACTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

Db 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAAATCATTTTGTCTGTGGCTCTGTGGCTTGACCATGGCCGCACTTCCCTG 2100
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGAGGCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db 2161 AGCACATGGGCTACATGCTGCTCATCTGCTCAATTCCTTTGCTTTCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATAATTGG 2280
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2281 GACTGCTCTATGTGTAACACATTTGCTCTTCCACCACTGCATCTCTAACTGC 2340
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2401 ATTAAGTTTATCTTCTGTGTGTAGTCCACTTCTCTGATGCTCAATCCCTTCTCTAC 2460
Qy 821 IleLeuPheAsnProHisPheLysGlyAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2461 ATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTGGAGCTTGAGAAAGCAACCTACGTC 2520
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2581 TCCTGTGACTCACTCAAGCTTGGTAACCTTTACCAGCTCCAGCATCACTATGACCTG 2640
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2641 CTCTCCAGTTCGCTGCCATCACCAGCTTATCCAGTACTGAGAGCTGCGCATCTTCTCTCT 2700
Qy 901 ValAlaPheValProCysLeu 907
Db 2701 GTGGCATTTGTCCCATGFTC 2721

RESULT 5

US-10-295-027-1114
; Sequence 1114, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: ~~Seal~~ Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666

1020
Seg 1380
Coln
par 68

Db 1669 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1728
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1729 AGAAATGGAGTGTGGACCATAGCAGTCTGGCACCTTACTTGTAACTTTGGTGACTTCA 1788
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleIysLeuLeuIleGlyValIleAla 600
Db 1789 ACAGTTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTAAATGGGGTCAATCGCA 1848
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1849 GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGTGTGTGGATGGCTTCACT 1908
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTTGGCAGCTTTTGACACGATGCTGCTGGTGGAGATGGGGTTGGTTGCCATGTCATT 1968
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheIleLeuThrIleAlaLeu 660
Db 1969 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTCTTCTGCTTACTTGGCAGCCCTG 2028
Qy 661 GluArgGlyPheSerValIysTyrSerAlaIysPheGluThrIysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC 2088
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
Qy 701 LeuGlyGlySerIysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGGTGGCAGCAAGTATGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuIleLeuLeuAsnSerLeuCysPheLeuMet 740
Db 2209 AGCACCATGGGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2268
Qy 741 ThrIleAlaTyrThrIysLeuTyrCysAsnLeuAspIysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATATTGG 2328
Qy 761 AspCysSerMetValIysHisIleAlaLeuLeuPheThrAsnCysIleIleLeuAsnCys 780
Db 2329 GACTGCTCTATGTGTAACACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2389 CCTGTGGCTTTCTGTGCTTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTGAAGTA 2448
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2449 ATTAAGTTTATCTTCTGTGTGTAGTCCCACTTCTGCTGATGCTCAATCCCTTCTCTAC 2508
Qy 821 IleLeuPheAsnProHisPheIysGluAspLeuValSerLeuArgIysGlnThrTyrVal 840
Db 2509 ATCTTGTTCATCTCCTTTAAGAGGATCTGCTGAGGCTCGAGAAAGCAAACTACGTC 2568
Qy 841 TrpThrArgSerIysHisProSerLeuMetSerIleAsnSerAspValGluIysGln 860
Db 2569 TGGACAAGATCAAAACCAACCAAGCTTGATGTCAATTAACCTTGTATGATGTGCAAAACAG 2628
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2629 TCTGTGATCTCAACTCAAGCTTGTGTAACCTTTTACAGGCTCCAGATCACTTATGACCTG 2688
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2689 CCTCCAGTTCCTGGCCATCACCCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2748
Qy 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCATTTTGTCCCATGTCTC 2769

RESULT 6

US-10-482-029-157
; Sequence 157, Application US/10482029
; Publication No. US20050037445A1

GENERAL INFORMATION:

; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 2880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-482-029-157

Alignment Scores:

Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-751-736-84 (1-907) x US-10-482-029-157 (1-2880)

Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 49 ATGGACACCTCCCGGCTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108
Qy 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 109 GGGGGCAGCTCTCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 169 GAGCCCGCAGCGGAGGATGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTGGAGCTG 228
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 229 CCTTCCAACTCAGCGTCTTCACTCTCCTACCTAGACCTCAGTATGAACAACTCAGTCAG 288
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
Db 289 CTGCTCCCGAATCCCTGCTGCCAGTCTCCGCTTCTGAGGAGGTAGCTCTTGGGGAAAC 348
Qy 101 AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMet 120
Db 349 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGGCTTTACAGTCTTAAAGTTCTTATG 408
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 409 CTCGAGAATAATCAGCTAAGACACGTACCCACAGAAGCTCTCGAGAATTTGGAGACCTT 468
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 469 CAATCCCTGCGTCTGGATGCTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 528
Qy 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluLeuProValGln 180
Db 529 CTGCATTTCCCTGAGCACCTGTGGCTGGATGACATGCGTTAACAGAAATCCCGCTCCAG 588
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIysIleHisHis 200
Db 589 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGCCCTTGAACAAAATACACCA 648
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn 220
Db 649 ATACCAGACTATGCTTTTGGAAACCTTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 708

NGD

Qy	221	ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp	240
Db	709	AGAATCCACTCCCTGGGAAAGAAATGCTTGTGATGGCTCCACAGCCTAGAGACTTTAGAT	768
Qy	241	LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrIleuSerAsnLeuLys	260
Db	769	TTAAATTTACAAATAACCTTGATGAATCCCACTGCAATTAGGACACTCTCCAACTTTAAA	828
Qy	261	GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn	280
Db	829	GAACCTAGATTTCAATAGCAACATATCAGGTGCTATCCTGAGAAGACATTTGTAGGCAC	888
Qy	281	ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla	300
Db	889	CCTTCTCTTATTACAATACATTTCTATGACAATCCCAATCCAAATTTGTGGGAGATCTGCT	948
Qy	301	PheGlnHisIleuProGluLeuArgThrIleuAsnGlyAlaSerGlnIleThrGlu	320
Db	949	TTTCAAACATTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATGAA	1008
Qy	321	PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrIleuThrGlyAlaGlnIle	340
Db	1009	TTTCTGATTTAACTGGAACTGCAAACTGGAGAGCTGTGACTTTAACTTGAGACACAGATC	1068
Qy	341	SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer	360
Db	1069	TCATCTCTCTCTCAACCGTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCT	1128
Qy	361	TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp	380
Db	1129	TACAACCTATTAGAAGATTTTACCACGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC	1188
Qy	381	LeuArgHisGlnGluIleTyrGluLysValAspThrPheGlnGlnLeuLeuSerLeu	400
Db	1189	CTAAGACATAATGAATCTAGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGGCCT	1248
Qy	401	ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr	420
Db	1249	CGATCGCTGAATTTGGCTTGGAACAAAATTGCTATTATTACCCCAATGCATTTTCCACT	1308
Qy	421	LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr	440
Db	1309	TTGGCATCCCTTAAATAAGCTGGACCTATCGTCCAAACCTCTCTGTCTGCTTTCTATAACT	1368
Qy	441	GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle	460
Db	1369	GGGTTTACATGGTTTAACTCACTTAAATTTAAACAGAAATCATGCTTTACAGAGCTTGATA	1428
Qy	461	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys	480
Db	1429	TCATCTGAAAACCTTTCCAGAACTCAAGGTATAGAAATGCCTTATGCTTACCAGTGTGT	1488
Qy	481	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
Db	1489	GCATTTGGAGTGTGTGAGATGCCCTATAAAGATTTCTAATCAATGGAATAAAGGTGACAC	1548
Qy	501	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
Db	1549	AGCAGTATGGACGACCTTCAATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1608
Qy	521	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540
Db	1609	GACCTTGNAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTCGAG	1668
Qy	541	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
Db	1669	TGTTTCACTTCCCCAGGCCCTTCAAACCCCTGTGAAACACCTGCTTGTATGGCTGGCTGATC	1728
Qy	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
Db	1729	AGAAATGGAGTGTGGACATAGCAGTTCTGGCACTTACTTGTGAATGCTTTGGTGACTTCA	1788
Qy	581	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600

Db	1789		ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACATGTTAATTGGGGTCATCGCA	1848
Qy	601	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620	
Db	1849	GCAGTGAACATGCTCAGGGAGTCTCCAGTCCGTGCTGGCTGGTGTGGATGCTTCACT	1908	
Qy	621	PheGlySerPheAlaArgHisGlyAlaTTrpTgLuAsnGlyValGlyCysHisValIle	640	
Db	1909	TTTGGCAGCTTTGCAACACATGGTGCCTGTGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT	1968	
Qy	641	GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu	660	
Db	1969	GGTTTTTGTGCCATTTTGGCTTCAGAAATCATCHTTTTCTGCTTACTCTGGCAGGCCCTG	2028	
Qy	661	GluArgGlyPheSerValIlySerAlaLysPheGluThrIlyLysAlaPropheSerSer	680	
Db	2029	GAGCGTGGGTTCCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2088	
Qy	681	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700	
Db	2089	CTGAAGATTAATCAATTTTGGCTCTGTGGCCCTGCTGGCCCTTGACCAGGCCGAGTTTCCCTG	2148	
Qy	701	LeuGlyGlySerIlySerGlyAlaSerProLeuCysLeuProLeuPropheGlyGluPro	720	
Db	2149	CTGGGTGGCAGCAGATATGCGCGCTCCCTCTCTGCTGCTCTTTTGGCGAGCCC	2208	
Qy	721	SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740	
Db	2209	AGCACCATGGGTACATGGTGCCTCATCTGCTCAATTCCTCTTGTCTTCCATCATGATG	2268	
Qy	741	ThrIleAlaTyrThrIlyLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr	760	
Db	2269	ACCATTCCTTACACCAAGCTCTACTGCATNTTGGACNAGGGAGACCTGGAGATATTTGG	2328	
Qy	761	AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780	
Db	2329	GACTGCTCTATGGTAAACACATTTGGCCCTGTTGCTCTTCAACCAACTGCATCTCAAACTGC	2388	
Qy	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800	
Db	2389	CCTGTGGCTTCTTGTCCCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCGAGTGA	2448	
Qy	801	IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr	820	
Db	2449	ATTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCTGTCATGTCATTCATCCCTTCTCTAC	2508	
Qy	821	IleLeuPheAsnProHisPheLysGlyAspLeuValSerLeuArgLysGlnThrTyrVal	840	
Db	2509	ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGGAGCTGTAGAAGCAAACTACGTC	2568	
Qy	841	TyrThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln	860	
Db	2569	TGGACAGATCAAAACACCCAGCTTGATGTCATTAACCTTCATGATGTCGAAAAACAG	2628	
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880	
Db	2629	TCCTGTGACTCAACTCAAGCGTTGGTAACCTTTACCAGCTCCAGACATCACTTATGACCTG	2688	
Qy	881	ProProSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900	
Db	2689	CCTCCCAGTTCCCGGCATCACACAGCTTATCCAGTGTACTGAGAGCTGCCATCTTCTCTCT	2748	
Qy	901	ValAlaPheValProCysLeu	907	
Db	2749	GTGGCATTTGTCCCATGTCTC	2769	

RESIT.T 7

US-10-651-237-48
; Sequence 48, Application US/10651237
; Publication No. US20050048494A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.


```
QY 621 PheGlySerPheAlaArgHisGlyAlaThrTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTTGGCAGCTTTGGCAGCAGAGTCCCTGGTGGGAGAAATGGGTGGTGGCAATGTCAAT 1968
QY 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1969 GGTTTTGTGTCATTTTGTCTCAGAAATCAATGTCTTCTGCTTACTCTGGCAGCCCTG 2028
QY 661 GluArgGlyPheSerValLysTySerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2088
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAAGTAATCATTTTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
QY 701 LeuGlyGlySerLysTySerGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGGTGGCAGCAGTATGGGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuMet 740
Db 2209 AGCACCATGGGCTACATGGTGGCTCTCATCTTGCTCAATTCCTTTGCTTCTCATGATG 2268
QY 741 ThrIleAlaTyThrLysLeuTyCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATGCTCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGGAATATTTGG 2328
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2329 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTTCTTCCCACTGATCTTAACTGTC 2388
QY 781 ProValAlaPheLeuSerPheSerSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 800
Db 2389 CTTGGGCTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2448
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyR 820
Db 2449 ATTAAGTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2508
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyRVal 840
Db 2509 ATCTGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2568
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
Db 2569 TGGACAAGATCAAAACACCCAGCTTGATGTCAATTAATCTCTGATGATGTGCAAAAACAG 2628
QY 861 SerCysAspSerThrGluAlaLeuValThrPheThrSerSerSerIleThrTyRAspLeu 880
Db 2629 TCCTGTGATCAACTCAAGCCCTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2688
QY 881 ProProSerSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer 900
Db 2689 CCTCCAGTTCCGTGCCATCACCACCTTATCCAGTACTGAGAGCTGCCATCTTCTCTCT 2748
QY 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCAATTTGCCCATGCTC 2769
```

RESULT 8

```
US-10-782-413-48
; Sequence 48, Application US/10782413
; Publication No. US20050048526A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: Colorectal Cancer Prognostics
; FILE REFERENCE: VDX-5002-GIP
; CURRENT APPLICATION NUMBER: US/10/782,413
; CURRENT FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: 10/551,237
```

```
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 2880
; TYPE: DNA
; ORGANISM: human
US-10-782-413-48
```

```
Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
```

US-10-751-736-84 (1-907) x US-10-782-413-48 (1-2880)

```
QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 49 ATGACACCTCCCGCTCGGTGCTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108
QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 109 GGGGGCAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCATTGC 168
QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 169 GAGCCGACGCGCAGATGTTGCTCAGGTGACCTGCTCCGACCTGGGGCTCTCGAGGCTG 228
QY 61 ProSerAsnLeuSerValPheThrSerTyLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 229 CCTTCCAACTCAGCGCTTTCACCTCTCTACCTAGACCTCAGTATGAAACAACATCAGT 288
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGlnLeuArgLeuAlaGlyAsn 100
Db 289 CTGCTCCCGAATCCCTGCCAGCTCTCGCTTCTCGAGGAGTTACGCTTCGCGGAAC 348
QY 101 AlaLeuThrTyRlleProLysGlyAlaPheThrGlyLeuTyRSerLeuLysValLeuMet 120
Db 349 GCTCTGACATACATTCACAGGGAGCATTCACCTGGCTTTTACAGTCTTAAAGTTCTTATG 408
QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 409 CTGCAGAAATAATAGCTAAGACACGTAACCAAGCTCTGCAGAAATTTGCGAAGCCTT 468
QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyRValProProSerCysPheSerGly 160
Db 469 CAATCCCTGCTCTGGATGCTAACCAACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGC 528
QY 161 LeuHisSerLeuArgHisLeuTyRLeuAspAsnAlaLeuThrGluIleProValGln 180
Db 529 CTGCATTCCTTGAGCAGCTGTGGCTGGATGCAATGCGTTAAACAGAAATCCCGTCCAG 588
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db 589 GCCTTTAGAGTTTATCGGCATTGCAAGCCATGACCTTGGCCCTGCAACAAATACACCAC 648
QY 201 IleProAspTyRAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 649 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 708
QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 709 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGTCCACAGCCTAGACACTTTAGAT 768
QY 241 LeuAsnTyRAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 769 TTAATATTACAATAACCTTGGATTCCTCCACTGCAATTAGGACACTCTCAACCTTAAA 828
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
```

Db 829 GAACAGGATTTTCATAGCAACAAATATACAGTCCGATACCTGAGAAAGCAATTTTAGGCAAC 888
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 889 CCTTCTCTTATTACAATAATTTCTATGACAATCCCATCCAATTTGTTGGAGAGATCTGCT 948
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 949 TTTCACCAATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCACAATAAATGAA 1008
Qy 321 PheProAspLeuThrGlyThrAlaAlaLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1009 TTTCCTGATTAACCTGGAACTGCAAACTGGAGAGCTGACTTTAACTTGGGACACAGATC 1068
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1069 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGCT 1128
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1129 TACAACCTTATTAGAAGATTTTACCAGTCTGCAAAAGCTTCAGAAAAATTGAC 1188
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1189 CTAAGACATATGAATCTACGAATTAAGTTGACACTTTCACAGAGTTGCTTAGCCTC 1248
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db 1249 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCTTTCCACT 1308
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1309 TTGGCATCCCTAATAAAGCTGACCTATGCTCAACCTCTGCTGCTTTTCTCTATAACT 1368
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1369 GGGTTACATGGTTAACTCATTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1428
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1429 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT 1488
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1489 GCATTTGGAGTGTGTGAGAATCCCTATAGATTTCTAATCAATGAATAAAGGTGACAAAC 1548
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1549 AGCAGTATGGAGACCTTCATAGAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1609 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1668
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleuIle 560
Db 1669 TGTTTCACTTCCCGAGGCCCTTCAAAACCTGTAACACCTGCTGATGCTGCTGCTGATC 1728
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1729 AGAATTTGAGTGTGACCATAGCAGTTCTGGCAGCTTACTTGTAAATGCTTTGTGTGACTTCA 1788
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1789 ACAGTTTTCAGATCCCTCTGATCATTTCCCCCATTAACCTGTTAAATTTGGGTCATCGCA 1848
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1849 GCAGTGAACATGCTCAGCGAGTCTCCAGTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTTGGCAGCTTTGGCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968

Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 1969 GGTTTTTTGTCCATTTTGTCTTCAATCATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2028
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC 2088
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAAGTAATCATTTTGTCTGTGCTCTGTGGCTGTGGCTTGGACCATGCGCGCAGTTCCCTCG 2148
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGTGGCAGCAAGTATGGCGCTCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGCC 2208
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2209 AGCACATGGGCTACATGCTGCTCTCATCTTGTCAATTCCTTTGCTTCTCTCATGATG 2268
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAATATTGG 2328
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2329 GACTGCTCTATGTGTAACACATTTGCTCTTCCCAACTGCATCTCTAAACTGC 2388
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2389 CCTGGGCTTCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2448
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2449 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2508
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2509 ATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2568
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
Db 2569 TGGCAAGATCAAAACACCCCAAGCTTGATGCAATTAATCTCTGATGATGCGAAAAAC 2628
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
Db 2629 TCCTGTGACTCACTCAAGCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2688
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2689 CTTCCAGTTCGCTGCCATCACCAGCTTATCCAGTGTAGAGCTGAGAGCTGCTTCTTCTCT 2748
Qy 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCATTTGCTCCCATGTCTC 2769

RESULT 9

US-10-295-027-848

; Sequence 848, Application US/10295027

; Publication No. US20030232350A1

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

Db 261 GGGGCGAGCTCTCCAGGCTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATTGC 320
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 321 GAGCCGACGGCAGGATGTGCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGAGCTG 380
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 381 CCTTCCAACTCAGCGTCTTACCTCTCCCTAGACCTCAGTATGACACATCATGTGAG 440
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuLeuArgLeuAlaGlyAsn 100
Db 441 CTGCTCCCGAATCCCTCCAGCTCTCCGCTTCTCTGGAGGATTACGCTTTGGGGAAAC 500
Qy 101 AlaLeuThrTyrIleProGlyGlyAlaPheThrGlyLeuTyrSerLeuLeuValLeuMet 120
Db 501 GCTCTGACATCATTTCCCAAGGGAGCATTCACCTGGCCCTTACAGTCTTTAAAGTCTTATG 560
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 561 CTGCAGAAATATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 620
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 621 CAATCCCTGCGTCTGGATGCTAAACCACATCAGCTATGTGCCCCCAAGCTGTTTTCAGTGGC 680
Qy 161 LeuHisSerLeuArgHisLeuTyrIleAspAspAsnAlaLeuThrGluIleProValGln 180
Db 681 CTGCATTCCTCGAGCACCTGCTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 740
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db 741 GCTTTTAAAGTTATTCGGCATTCGACCATGACCATGACCTTGGCCCTGGAACAAATACACAC 800
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 801 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 860
Qy 221 ArgIleHisSerLeuGlyLysIleCysPheAspGlyLeuHisSerLeuLeuThrLeuAsp 240
Db 861 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGACACTTTAGAT 920
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 921 TTAATTAACAATAACCTTGATGAATTCGCCACTGCCAATTAGACACTCTCCAACTTAAA 980
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 981 GAACTAGGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 1040
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1041 CCTTCTCTTATTACAATAATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 1100
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1101 TTTTCAACATTTACCTGAACATAGAACACTGACTCTGAATGTGTCCTCAAAATAACTGAA 1160
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1161 TTTTCTGATTTAACTGGAACTGCAAACTCGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1221 TCATCTCTTCTCTCAAAACGCTCGCAATCAGTTACCTAATCTCCAAAGTCTAGATCTGCT 1280
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1281 TACAACCTATTAGAAGATTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAATATGAC 1340
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1341 CTAAACATATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1400

Qy 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1401 CGATCGCTGAATTTGGCTTGGACAAATTCCTATTATTTCACCCCAATGCATTTTCCACT 1460
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1461 TTGCCATCCCTAATAAAGCTGACCTATCGTCCAAACCTCCTGTCTCTTTTCCATAACT 1520
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1521 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTTACAGAGCTTGATA 1580
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1581 TCATCTGAAACTTTTCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACAGTCTGT 1640
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTyrAsnLysGlyAspAsn 500
Db 1641 GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGAATAAAGGTGACAAC 1700
Qy 501 SerSerMetAspAspLeuHisLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1701 AGCAGTATGGACGACCTTCATAAGAAAGATCTCGAAATGTTTTCAGGCTCAAGATGAACGT 1760
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1761 GACCTTGAAGATTTCTGCTTGACTTTGAGAGAAAGCTGAAAGCCCTTCATTCAGTGGAG 1820
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Db 1821 TGTTCACTTCCCCAGGCCCTTCAAACTCTGTAACACCTGCTGTGATGGCTGGCTGATC 1880
Qy 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1881 AGAATGGAGTGTGGACCATAGCAGTTCTGCACTTACTTGTAAATGCTTTTGGTCACTTCA 1940
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuLeuGlyValIleAla 600
Db 1941 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACGTAAATTTGGGGTCAATCGCA 2000
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2001 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGGCTGTGTGATGGCTGCTCACT 2060
Qy 621 PheGlySerPheAlaArgHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIle 640
Db 2061 TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGGTGGTTCATGTCAAT 2120
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 2121 GGTTTTTTGTCATTTTGTCTTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2180
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2181 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2240
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2241 CTGAAGTAATCAATTTTGTCTGTGCTGCTGCTGCTTGCATGCGCGCAGTTCCCTCTG 2300
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2301 CTGGGTGGCAGCAAGATGGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGAGGCC 2360
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2361 AGCACCATGGGCTACATGGTCTCATCTGTCTCAATTCCTTGTCTTCTCTCATGTATG 2420
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
Db 2421 ACCATTGCTACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTCGAGAAATATTTGG 2480

Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAACAAAATTGCTATTATTACCCCAATGCAATTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCGTGTCTTTTCCCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGAAATCATGCCCTACAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProfyrAlaIleGlnCysCys 480
Db 1381 TCATCTGAAATCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrlsIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCCATTAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGACGACCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAGATTTCCCTGCTTGACTTTTGAGAGACCTGAAAGCCCTTCATTCACTGAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisIleLeuLeuAspGlyTrpLeuIle 560
Db 1621 TGTTTCACTTCCCCAGGCCCTTCAAACCCCTGTGAACACCTGCTGTGATGGCTGCTGCATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATTTGGAGTGTGGACATAGCAGTCTTGCACTTTACTTGTATGCTTTGGTGACTTCA 1740
Qy 581 ThrValPheArgSerProLeuTyrlsSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACGTGTAATTTGGGGTCATGCCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCGCTGCTGGTGTGGATGGCTTCCACT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGCACACATATGGCTCTGGTGGGAGAAATGGGTTGGTTCCTATGTCATT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTGTGTCATTTTGTCTTCAGATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrlsSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAAATCAATTTGGCTCTGTGCCCTGTGGCCCTTGACCATGGCCGAGTTCCTCGT 2100
Qy 701 LeuGlyGlySerLysTyrlsGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGTATGGGCCCTCCCTCTCTGCTTGCCTTGGCTTTTGGGAGCCC 2160
Qy 721 SerThrMetGlyTyrlsMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATCCCTTGTCTTCTTCTCATGATG 2220
Qy 741 ThrIleAlaTyrlsLeuTyrlsCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTCCCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATATTGG 2280
Qy 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780

Db 2281 GACTGCTCTATGAA-AAAAACACATTGCCCTGTGCTCTTCCACCAACTGSCATCCTAAACTG 2339
Qy 780 sProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
Db 2340 CCCTGTGGCTTCTCTGTGCTTCTCTCTTTAAATAAACCTTACATTTATCAGTCTGAAGT 2399
Qy 800 lIleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
Db 2400 AATTAAGTTTATCTTCTGTGTGTAGTCCACATCTCTGCATGCTCAATCCCTTCTCTA 2459
Qy 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyVa 840
Db 2460 CATCTGTTCAATCCTCACCTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAACCTACGT 2519
Qy 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysG1 860
Db 2520 CTGGACAAGATCAAAACACCAAGCTTGATGTCAATTAACCTCGATGATGTGCAAAAACA 2579
Qy 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyAspLe 880
Db 2580 GTCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCT 2639
Qy 880 uProProSerSerValProSerProAlaTyrlsProValThrGluSerCysHisLeuSerSe 900
Db 2640 GCCTCCAGTTCGGTGCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCCTC 2699
Qy 900 rValAlaPheValProCysLeu 907
Db 2700 TGTGGCATTTGTCCCATGTCTC 2721

RESULT 12
US-10-505-486-205
; Sequence 205, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 205
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Human
; US-10-505-486-205

Alignment Scores:
Pred. No.: 0 Length: 3438
Score: 806.00 Matches: 906
Percent Similarity: 99.78% Conservatives: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 88.86% Indels: 2
DB: 21 Gaps: 0

US-10-751-736-84 (1-907) x US-10-505-486-205 (1-3438)

Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 1 ATGACACACCTCCCGGCTCGGTGTGCTCTCTCTGCTGTGCTGTGCTGAGTGGCGACC 60
Qy 21 GlyClySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 61 GGGGGCAGCTCTCCAGGCTGTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCATGC 120
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60

Db	121	GAGCCCGACGCGACGAGTGTGGCTCAGGGTGGACCTGCTCCGACCTGGGGGCTCTCGAGCGTG	180
Qy	61	ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln	80
Db	181	CGTTCCAAACCTCAGCGTCTCACCTCTACTACAGCCTCAGTAGAACACATCAGTCAG	240
Qy	81	LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn	100
Db	241	CTGCTCCGAATCCCTGCCAGTCTCCGCTCTCTGGAGGAGTTACGCTTGGCGGAAC	300
Qy	101	AlaLeuThrTyrIle-ProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMe	120
Db	301	GCTCTGCATACGTTCCCAAGGAGGACATTCACCTGGGCTTTACAGTCTTAAAGTTCAT	359
Qy	120	LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLe	140
Db	360	GCTGCAGATTAATCAGCTTAAGACAGCTACCCACAGAGAGCTCTGCAGAAATTCGAGCGCT	419
Qy	140	uGlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGln	160
Db	420	TCAACTCCCTGGCTGGATGTAACCAACATCAGCTATGTGCCCAAGCTGTTCAGTGG	479
Qy	160	YLeuHisSerLeuArgHisIleuTrpLeuAspAsnAlaLeuThrGluIleProValGln	180
Db	480	CGTGCATTCCTTGAGGACCTGTGGCTGGATGACAATCGCTTAACAGAAATCCCGCTCCA	539
Qy	180	nAlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIysIleHisI	200
Db	540	GGCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAACAAATACACCA	599
Qy	200	sIleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAs	220
Db	600	CATACACAGACTATGCCTTTGGAAACCTCTCCAGCTTGTAGTTCTACATCTCCATAACAA	659
Qy	220	nArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAs	240
Db	660	TAGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGA	719
Qy	240	pLeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLy	260
Db	720	TTTAAATTTACAAATAACTTTGATGAATTTCCCCACCTGCAATTAGGAACCTCTCCAACCTTAA	779
Qy	260	sGluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAs	280
Db	780	AGAACTAGGATTTTCATAGCAACAATATCAGGTGCTATACCTGAGAAAGCATTTGTAGGCAA	839
Qy	280	nProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAl	300
Db	840	CCCTTCTCTATTATCAATATACATTTCTATGACAATCCCATCCCAATTTGTTGGAGACTCGC	899
Qy	300	aPheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGln	320
Db	900	TTTTCAACATTTACTGAACTTAAGAACACTGACTGTAATGGTGCCCTCACAAATTAACCTGA	959
Qy	320	uPheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle	340
Db	960	ATTTCTCTGATTTAACTGGAAGCTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGAT	1019
Qy	340	eSerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSe	360
Db	1020	CTCATCTCTCTCTCAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT	1079
Qy	360	rTyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnIysLeuGlnIysIleAs	380
Db	1080	TTACAACTATTTAGAAGATTTTACCACGATTTTTCAGTCTGCCAAAGCTTCAGAAATTTGA	1139
Qy	380	pLeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLe	400
Db	1140	CCTAAGACATTAATGAAATCTACGAAATTAAGATTGACACTTTCACGACGATTGCTTAGCCT	1199
Qy	400	uArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr	420

1200	CCGATCGCTGAA	TTTGGCTTGGAA	CAAAAATTGCTATTATTTACCCCAATGCATTTTCCAC	1259
420	rLeuProSerLeu	lleLysLeuAspLeuSerSerAsnLeuSerSerPheProIle	rH	440
1260	TTTGCCATCCCT	ATAAAGCTGGACCTATCGTCCAACCTCTCTGCTGCTCTTTTCCATAAC	1319	
440	rGlyLeuHisGlyLeu	ThrHisIleLysLysLeuThrGlyAsnHisAlaLeuGlnSerLeu	1460	
1320	TGGGTTACATGG	TTTAACCTCACTTAAATAATTAACAGGAAATCATGCCTTACAGAGCTTGAT	1379	
460	eSerSerGluAsn	PheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCy	480	
1380	ATCATCTGAA	AAACTTTCCAGACTCAAGGTATAGAAATGCCCTATGCTTTACCAGTGTCTG	1439	
480	sAlaPheGlyVal	CysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAs	500	
1440	TGCATTTGGAG	TGTGAGAAATGCCATAAGATTTCTTAATCAATGGAATAAAGGTGACAA	1499	
500	nSerSerMetAsp	AspLeuHisIlyLysLysAspAlaGlyMetPheGlnAlaGlnAspGluAr	520	
1500	CAGCAGTATG	GCACGCTTCATAGAAGAATGCTGGAAATGTTTCAGGCTCAAGATGAACG	1559	
520	qAspLeuGluAsp	PheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValG	540	
1560	TGACCTTTGA	AGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCA	1619	
540	nCysSerProSer	ProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeu	560	
1620	GTGTTACCT	TTCCCCAGGCCCTTCAAACTCTGGAACACCTGCTTGATGGCTGGCTGAT	1679	
560	eArgIleGlyVal	TrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSe	580	
1680	CAGAAATGG	AGTGTGGACCATAGCAGTTCCTGGCACCTTACTTGTAAATGCTTTGTGTGACTTC	1739	
580	rThrValPheArg	SerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAl	600	
1740	AACAGTTTT	CAGATCCCTCTGACATTTCCCCCATTTAAACTGTTAATGGGGTCATCGC	1799	
600	aAlaValAsnMet	LeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheTh	620	
1800	AGCAGTGAA	CATGCTCACGGGAGTCTCCAGTGGCGTCTGGCTGGTGGATGGGTTCAC	1859	
620	rPheGlySerPhe	AlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValII	640	
1860	TTTTGGCAG	CTTTGCACGACATGGTGCTGCTGGTGGGAATGGGGTGGTTGCCATGTTCAT	1919	
640	eGlyPheLeuSer	IlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLale	660	
1920	TGGTTTTTT	TGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTCGCTTACTCTGGCAGCCCT	1979	
660	uGluArgGlyPhe	SerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSe	680	
1980	GGAGCGTGG	GTCTCTGTGAAATATCTGCBAATTTGAAACGAAAGCTCCATTTTCTAG	2039	
680	rLeuLysValIle	IleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLe	700	
2040	CCTGAAAGT	TAATCATTTTGTCTGTGCCCTGTCTGGCTTCACTGCTTACTCTGGCAGTTC	2099	
700	uLeuGlyGlySer	IlySerIlyGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPr	720	
2100	GCTGGGTGG	CAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGGGAGCC	2159	
720	oSerThrMetGly	TyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMe	740	
2160	CAGCACCAT	GGGCTACATGCTCGCTCTCATCTTGCTCAATTCCTTTGCTTCTCTCATGAT	2219	
740	tThrIleAlaTyr	ThrLysLeuTyrCysAsnLeuLeuAspLysGlyAspLeuGluAsnIleTr	760	
2220	GACCATTTGC	TTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAATATT	2279	
760	pAspCysSerMet	ValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy	780	
2280	GGACTGCTAT	TGTTAAACACATGCGCTTGTGCTCTTCACCACTGCATCTCTAAACTG	2339	

Db 1627 ATAGCAGTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCAACAGTTTTTCAGATCCCT 1686
Qy LeuTyrIleSerProIleGlyLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 606
Db CTGTACATTTCCCCATTAAACCTGTTAATTTGGGTTCATCGCAGCAGTGAACATGCTCAG 1746
Qy GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 626
Db GGAGTCTCCAGTCCGCTGCTGCTGGTGGATGCGTTCACTTTTGGCAGCTTTGCACGA 1806
Qy HisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe 646
Db CATGGTCTGGTGGAGATGGGGTGGTGGCCATGTCATTTGGTTTGTCCATTTT 1866
Qy AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal 666
Db GCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGTCTCTGTG 1926
Qy LysTyrSerAlaLysPheGluThrIlylsAlaProPheSerSerLeuIlylsValIleIleLeu 686
Db AAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTGAAAGTAAATCATTTTG 1986
Qy LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerIlylsTyr 706
Db CTCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
Qy GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 726
Db GCGGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106
Qy VallalaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrIlyls 746
Db GTGCTCTCTCATCTGCTCAATTCCTTTCCTTCTCATGATGACCATTTGCCCTACACCAAG 2166
Qy LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyrAspCysSerMetValIlyls 766
Db CTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGGAGCTGCTCTATGTAATA 2226
Qy HisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer 786
Db CACATTGCTGCTGCTGCTCTTCCACCACTGCATCTTAACCTGCCCTGTGCTGCTTCTTGCC 2286
Qy PheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeu 806
Db TTCTCTCTTTAAATAAACCTTACATTTATCAGTCTGGAAGTAAATTAAGTTTATCTCTG 2346
Qy ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis 826
Db GTGGTAGTCCCACTCTCCGCAAGTCTCAATCCCTTCTCTACATCTTGTTCATCCCTCAC 2406
Qy PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTyrThrArgSerIlyshis 846
Db TTTAAGGAGGATCTGTGAGCTGGAAGCAACCTACCTGCTGACCAAGATCAAAACAC 2466
Qy ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln 866
Db CCAAGCTTGATGTCATTAATCTCATGATGTCGAAAAACAGTCTGTGACTCAACTCAA 2526
Qy AlaLeuValThrPheThrSerSerIleThrTyrAspLeuProProSerSerValPro 886
Db GCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTGCCCTCCAGTTCGGTGCA 2586
Qy SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 906
Db TCACAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCCATGT 2646
Qy Leu 907
Db 2647 CTT 2649
RESULT 14
US-10-295-027-1113
; Sequence 1113, Application US/10295027

Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1113
LENGTH: 2651
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-1113
Alignment Scores:
Pred. No.: 0 Length: 2651
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.47% Indels: 0
DB: 17 Gaps: 0
US-10-751-736-84 (1-907) x US-10-295-027-1113 (1-2651)
Qy 287 HisPheTyrAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 306
Db 787 CATTTCTATGACATCCCATCCATTTGTTGGAGATCTGCTTTCAACATTTACTGAA 846
Qy 307 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly 326
Db 847 CTAAGAACACTGACTCTGAATGGTGCCTCAAAAATACTGAATTTCTCTGATTTAACTGGA 906
Qy 327 ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 346
Db 907 ACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCAGAGATCTCATCTCTTCTCCCAACC 966
Qy 347 ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAsp 366
Db 967 GTCTGCAATCAGTTACTTCAATCTCAAGTGTAGATCTGTCTTACACCTATTAGAGAT 1026
Qy 367 LeuProSerPheSerValCysGlnLysLeuGlnLysIleAspLeuArgHisAsnGluIle 386

1027	Db		TTACCCAGT	TTTT	CAGT	CTGC	CAAA	AAG	CTT	CAG	AAAA	TT	GAC	CTT	TA	GAC	ATA	T	GAA	ATC	1086	
387	Qy		TyrGlu	leu	lys	Val	Asp	Thr	Phe	Gln	Leu	Leu	Ser	Leu	Arg	Ser	Leu	Asn	Leu	Ala	406	
1087	Db		TACG	AAAT	TAA	AGT	TG	CAC	TTT	CC	G	CAG	T	GT	CTT	AG	CTC	CG	AT	T	1146	
407	Qy		TrpAsn	lys	Val	Ala	Ile	His	Pro	Asn	Ala	Phe	Ser	Thr	Leu	Pro	Ser	Leu	Ile	Lys	426	
1147	Db		TGG	AA	CA	AA	AT	TG	CT	AT	T	T	CA	CC	C	A	TG	CA	T	T	1206	
427	Qy		LeuAsp	Leu	Ser	Ser	Asn	Leu	Ser	Ser	Phe	Pro	Ile	Thr	Gly	Leu	His	Gly	Leu	Thr	446	
1207	Db		CTG	GA	CT	AT	CG	CA	CC	T	C	T	G	T	G	T	T	T	C	T	1266	
447	Qy		His	Leu	lys	Leu	Thr	Gly	Asn	His	Ala	Leu	Gln	Ser	Leu	Leu	Ser	Ser	Glu	Asn	466	
1267	Db		CAC	T	T	A	A	A	T	T	A	A	C	G	G	A	A	T	C	A	T	1326
467	Qy		Glu	Leu	lys	Val	Ile	Glu	Met	Pro	Tyr	Ala	Tyr	Gln	Cys	Ala	Phe	Gly	Val	Cys	486	
1327	Db		GA	CT	CA	AG	GT	T	A	G	A	A	T	G	C	T	T	A	C	A	T	1386
487	Qy		Asn	Ala	Tyr	lys	Val	Ser	Asn	Gln	Trp	Asn	lys	Gly	Asp	Asn	Ser	Ser	Met	Asp	506	
1387	Db		AAT	GC	T	A	T	A	A	G	T	T	C	A	T	CA	T	CA	T	AA	T	1446
507	Qy		His	lys	lys	Asp	Ala	Gly	Met	Phe	Gln	Ala	Gln	Asp	Glu	Arg	Asp	Leu	Glu	Asp	526	
1447	Db		CAT	A	A	G	A	A	A	G	A	T	G	T	T	T	T	T	T	T	T	1506
527	Qy		LeuAsp	Phe	Glu	Glu	Asp	Leu	lys	Ala	Leu	His	Ser	Val	Gln	Cys	Ser	Pro	Ser	Pro	546	
1507	Db		CTT	G	AC	T	T	T	G	A	G	A	CC	T	T	CA	T	CA	T	G	A	1566
547	Qy		Pro	Phe	lys	Pro	Cys	Glu	His	Leu	Leu	Asp	Gly	Trp	Leu	Ile	Arg	Ile	Gly	Val	566	
1567	Db		CC	T	T	CA	A	CC	T	G	T	GA	CA	CC	T	G	T	G	G	T	G	1626
567	Qy		Ile	Ala	Val	Leu	Ala	Leu	Thr	Cys	Asn	Ala	Leu	Val	Thr	Ser	Thr	Val	Phe	Arg	586	
1627	Db		AT	A	G	C	A	G	T	CT	T	CA	T	T	G	T	G	A	CT	T	T	1686
587	Qy		Leu	Tyr	Ile	Ser	Pro	Ile	lys	Leu	Leu	Ile	Gly	Val	Ile	Ala	Val	Asn	Met	Leu	606	
1687	Db		CT	G	A	CA	T	T	CC	CC	CA	T	T	AA	T	G	G	G	T	CA	T	1746
607	Qy		Gly	Val	Ser	Ser	Ala	Val	Leu	Ala	Gly	Val	Asp	Ala	Phe	Thr	Phe	Gly	Ser	Phe	626	
1747	Db		GG	A	G	T	CT	CA	G	T	G	C	G	T	G	T	G	A	T	G	A	1806
627	Qy		His	Gly	Ala	Trp	Trp	Glu	Asn	Gly	Val	Gly	Cys	His	Val	Ile	Gly	Phe	Leu	Ser	646	
1807	Db		CAT	G	T	G	T	G	C	T	G	G	A	T	G	G	G	T	T	G	T	1866
647	Qy		Ala	Ser	Glu	Ser	Ser	Val	Phe	Leu	Leu	Thr	Leu	Ala	Ala	Leu	Glu	Arg	Gly	Phe	666	
1867	Db		GC	T	T	C	A	GA	A	T	CA	T	G	T	T	T	T	T	T	T	T	1926
667	Qy		Lys	Tyr	Ser	Ala	Lys	Phe	Glu	Thr	Lys	Ala	Pro	Phe	Ser	Ser	Leu	Lys	Val	Ile	686	
1927	Db		AA	A	T	T	T	C	G	A	A	A	T	T	G	A	A	A	T	CA	T	1986
687	Qy		Leu	Cys	Ala	Leu	Leu	Ala	Leu	Thr	Met	Ala	Val	Pro	Leu	Leu	Gly	Gly	Ser	Lys	706	
1987	Db		CT	C	T	G	T	G	C	C	T	T	G	A	C	A	T	G	C	G	A	2046
707	Qy		Gly	Ala	Ser	Pro	Leu	Cys	Leu	Pro	Phe	Gly	Glu	Pro	Ser	Thr	Met	Gly	Tyr	Met	726	
2047	Db		GG	C	CC	T	CC	C	T	T	T	G	C	T	T	T	T	G	G	G	A	2106
727	Qy		Val	Ala	Leu	Ile	Leu	Leu	Asn	Ser	Leu	Cys	Phe	Leu	Met	Met	Thr	Ile	Ala	Tyr	746	

Db	2107	GTCCGCTCTCAATCTTGCTCAATCCCTTTTGCTTCTCATGATGACCAATGTCCTACACCAAG	2166
Qy	747	LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTirAspCysSerMetValIys	766
Db	2167	CTCTACTGCAATTTGGACAGAGGAGACCTGGAGATATTTGGGACTGCTCTATGGTAAAA	2226
Qy	767	HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer	786
Db	2227	CACATTGCCCTGTTGCTCTTTCACCAACTGCATCCTAAACTGCCCTGTGGCTTTCTTGTC	2286
Qy	787	PhaSerSerLeuIleAsnLeuThrPheIleSerProGluValIleIleLysPheIleLeuLeu	806
Db	2287	TTCTCCTCTTTATAAACCCTTACATTTATCAGTCCTGAAGTAATTAAGTTTATCCTTCTG	2346
Qy	807	ValIleValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis	826
Db	2347	GTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCAATCCTCAC	2406
Qy	827	PhelysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis	846
Db	2407	TTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAAACTACGCTCTGGACAAGATCAAAACAC	2466
Qy	847	ProSerLeuMetSerIleAsnSerAspAspValGluLysGlnSerCysAspSerThrGln	866
Db	2467	CCAAGCTTGATGTCATTAATCTCTGATGATGTGCAAAAACAGTCCTGTGACTCAACTCAA	2526
Qy	867	AlaLeuValThrPheThrSerSerSerIleThrTyrAspLeuProProSerSerValPro	886
Db	2527	GCCTTGTTGAACCTTTTACCAGCTCCAGATCACATTAATGACCTGCCTCCAGTTTCGGGCCA	2586
Qy	887	SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys	906
Db	2587	TCACCAGCTTATCCAGTGACTGAGAGCTGCCACTTTTCTCTGTGGCATTTGTCCCATGT	2646
Qy	907	Leu 907	
Db	2647	CTT 2649	
RESULT 15			
US-10-173-999-27			
; Sequence 27, Application US/10173999			
; Publication No. US20040005563A1			
; GENERAL INFORMATION:			
; APPLICANT: Mack, David H.			
; APPLICANT: Gish, Kurt C.			
; APPLICANT: Eos Biotechnology, Inc.			
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions			
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian			
; TITLE OF INVENTION: Cancer			
; FILE REFERENCE: 018501-002420US			
; CURRENT APPLICATION NUMBER: US/10/173,999			
; CURRENT FILING DATE: 2002-06-17			
; PRIOR APPLICATION NUMBER: US 60/299,234			
; PRIOR FILING DATE: 2001-06-18			
; PRIOR APPLICATION NUMBER: US 60/315,287			
; PRIOR FILING DATE: 2001-08-27			
; PRIOR APPLICATION NUMBER: US 60/350,666			
; PRIOR FILING DATE: 2001-11-13			
; PRIOR APPLICATION NUMBER: US 60/372,246			
; PRIOR FILING DATE: 2001-04-12			
; NUMBER OF SEQ ID NOS: 163			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 27			
; LENGTH: 2651			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-173-999-27			
Alignment Scores:			
Pred. No.:		0	Length: 2651
Score:		621.00	Matches: 621
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0

Query Match:	68.47%	Indels:	0
DB:	17	Gaps:	0
US-10-751-736-84 (1-907) x US-10-173-999-27 (1-2651)			
QY	287	HisPheTyrAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu	306
DB	787	CATTCTATGACAAATCCATCCAAATTTGTTGGGAGACTCTCTTTCAACATTTACCTGAA	846
QY	307	LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly	326
DB	847	CTAAGAACACTGACTCTGAATGGTCCCTCACAAATAACTGAAATTTCTGATTTAACTGGA	906
QY	327	ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr	346
DB	907	ACTGCAAACTCGGAGAGTCTGACTTTAACTGGAGCAGATCTCATCTCTCTCTCAAAAC	966
QY	347	ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAsp	366
DB	967	GTCTGCAATCAGTTACCTAATCTCCAAAGTGTAGATCTGTCTTACAACTATTAGAAGAT	1026
QY	367	LeuProSerPheSerValCysGlnLeuGlnLysIleAspLeuArgHisAsnGluIle	386
DB	1027	TTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAATTCACCTTAGACATATATGAATC	1086
QY	387	TyrGluIleLysValAspThrPheGlnGlnLeuSerLeuArgSerLeuAsnLeuAla	406
DB	1087	TACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCTCCGATCGCTGAATTTGGCT	1146
QY	407	TyrAsnLysIleAlaIleHisProAsnAlaPheSerThrLeuProSerLeuIleLys	426
DB	1147	TGGAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACTTTGCCATCCCTAAATAAG	1206
QY	427	LeuAspLeuSerSerAsnLeuSerSerPheProIleThrGlyLeuHisGlyLeuThr	446
DB	1207	CTGGACCTATCGTCCAACTCTCTGCTCTTTTCTTATTAATCTGGGTACATGGTTAACT	1266
QY	447	HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro	466
DB	1267	CACCTAAATTAACAGAAATCATGCTTACAGAGCTTGATATCATCTGAANAATTTCCA	1326
QY	467	GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu	486
DB	1327	GAACTCAAGGTATAGAAATGCTCTATGCTTACAGTCTGTGCAATTTGGAGTGTGAG	1386
QY	487	AsnAlaTyrLysIleSerAsnGlnTyrAsnLysGlyAspAsnSerSerMetAspAspLeu	506
DB	1387	AATGCCATTAAGATTTCTAATCAATGGAAATTAAGGTGACCAACAGCAGTATGGACGCTT	1446
QY	507	HisLysIysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu	526
DB	1447	CATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACGTGACCTTGAAGATTTCTG	1506
QY	527	LeuAspPheGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly	546
DB	1507	CTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGAGTGTTCACCTTCCCCAGGC	1566
QY	547	ProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIleArgIleGlyValTyrThr	566
DB	1567	CCCTTTCAAAACCTGTGAACACCTCTCTGATGGCTGGCTGATCAGAAATGGAGTGTGACC	1626
QY	567	IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro	586
DB	1627	ATAGCAGTTCTGGCAGCTTACTTGTAAATGCTTTGGTGAATTCACAGTTTTCAGATCCCT	1686
QY	587	LeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaValAsnMetLeuThr	606
DB	1687	CTGTACATTTCCCCCAATTAACCTGTTAATTTGGGTCTATCGCAGCAGTGAACATGCTCACG	1746
QY	607	GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg	626
DB	1747	GGAGTCTCCAGTCCGCTGGCTGGTGGTGGATCGCTTCACTTTTGGCAGCTTTGGCACGA	1806

Search completed: July 12, 2005, 13:53:20
Job time : 3118 secs

QY	627	HisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe	646
DB	1807	CATGGTGCCTGGTGGGAGAAATGGGGTGGTGCATGTCTATGGTTTTTTTGTCCATTTTT	1866
QY	647	AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal	666
DB	1867	GCTTCAGATCATCTGTTTCTCTGCTTACTCTGGAGCCCTGGAGCGTGGGTCTCTGTG	1926
QY	667	LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeu	686
DB	1927	AAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTGAAAGTAATCATTTTG	1986
QY	687	LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr	706
DB	1987	CTCTGTCCTCTGCTGGCTTGACCATGGCGCGAGTTCCCTGTCTGGGTGGCAGCAAGTAT	2046
QY	707	GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet	726
DB	2047	GGGCGCTCCCTCTCTGCTGGCTTTGCCCTTTTGGGAGCCAGACCATGGGCTACATG	2106
QY	727	ValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys	746
DB	2107	GTGCTCTCATCTTGTCAATTCCTTGTCTCTCATGATGACCATTTGCTACACCAAG	2166
QY	747	LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyrPheCysSerMetValLys	766
DB	2167	CTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTTGGGACTGCTCTATGGTAAAA	2226
QY	767	HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer	786
DB	2227	CATTTGCCCTGTGCTCTTCCAACTGCATGCACTTAACCTGCCCTGTGGCTTTCTGTCC	2286
QY	787	PheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeu	806
DB	2287	TTCTCTCTTTAATAAACCTTACATTTATCAGTCTGAAAGTAATTAAGTTTATCTTCTG	2346
QY	807	ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis	826
DB	2347	GTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTACATCTTTGTTCAATCTCCAC	2406
QY	827	PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTyrArgSerLysHis	846
DB	2407	TTTAAGAGGATCTGGTAGCCTGAGAAAGCAACCTACGCTCTGGACNAGATCAAAACAC	2466
QY	847	ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln	866
DB	2467	CCAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACAGTCTGTGACTCAACTCAA	2526
QY	867	AlaLeuValThrPheThrSerSerIleThrTyrAspLeuProProSerSerValPro	886
DB	2527	GCCTTGGTAACCTTTTACCAGCTCCAGCATCACATTATGACCTGCCCTCCAGTTCGGTGCCA	2586
QY	887	SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys	906
DB	2587	TCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTGTGGCATTTTGCCCATGT	2646
QY	907	Leu 907	
DB	2647	CTT 2649	

This Page Blank (uspto)

GenCore version 5.1.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:53:52 ; Search time 6975 Seconds
(without alignments)
4949.719 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTRRLGVLLSLPVLQLAT.....PAYPTVTSCHLSVAFVPCFL 907

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US1071736/runat.12072005.083641.22183/app.query.fasta.1.1095
-DB=EST -QFMT=fastap -SURFIX=oli.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1071736 @CGN.1.1.5210 @runat.12072005.083641.22183 -NCPV=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl1:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	764	84.2	2724	9 AY411732	Homo sapi
2	316	34.8	1060	1 AL541959	AL541959
3	235	25.9	800	4 BG196708	BG196708
4	179	19.7	590	6 CD617354	CD617354
5	176	19.4	873	5 BX403219	BX403219
6	158	17.4	475	5 BX107244	BX107244
7	157	17.3	1104	1 AL530798	AL530798
8	156	17.2	472	1 AA460529	AA460529
9	156	17.2	556	5 BQ323949	BQ323949

10	149	16.4	671	4	BM795259
11	137	15.1	634	7	CF177761
12	122	13.5	374	1	AI367723
13	118	13.0	358	1	AI699236
14	111	12.2	539	1	AI697103
15	98	10.8	2593	9	AY411733
16	95	10.5	494	1	AA424098
17	91	10.0	576	5	BX499910
18	83	9.2	276	6	CD617346
19	83	9.2	276	6	CD617347
20	83	9.2	276	6	CD617348
21	83	9.2	276	6	CD617349
22	83	9.2	276	6	CD617353
23	81	8.9	268	6	CD617342
24	81	8.9	271	6	CD617352
25	78	8.6	250	6	CD617356
26	74	8.2	222	6	CD617357
27	74	8.2	222	6	CD617358
28	72	7.9	243	6	CD617343
29	65	7.2	940	7	CK449409
30	64	7.1	330	5	BX357656
31	58	6.4	224	6	CD617359
32	57	6.3	432	8	AQ076504
33	56	6.2	275	6	CD617344
34	51	5.6	627	2	BB613285
35	51	5.6	763	7	CF742674
36	51	5.6	2724	9	AY411734
37	51	5.6	4796	3	AK047873
38	48	5.3	492	6	CD617355
39	46	5.1	397	1	AI539105
40	45	5.0	275	6	CD617350
41	45	5.0	275	6	CD617351
42	42	4.6	698	5	BF143829
43	41	4.5	542	1	AI962439
44	39	4.3	280	6	CD617345
45	39	4.3	396	1	AL119316

ALIGNMENTS

RESULT 1	AY411732	2724 bp	DNA	linear	GSS 16-DEC-2003
LOCUS	AY411732	Homo sapiens GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
DEFINITION	AY411732	AY411732.1	GI:39767700		
ACCESSION	AY411732				
VERSION	AY411732.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2724)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2724)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	location/Qualifiers				
source	1. .2724				


```

RESULT 3
BG196708
LOCUS RST15934 Athersys RAGE Library Homo sapiens cDNA, mRNA linear EST 21-APR-2001
ACCESSION BG196708
VERSION BG196708.1 GI:13718395
EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozer,S.,
Maye,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave., Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 481.
FEATURES
source
1..800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
ORIGIN
Alignment Scores:
Pred. No.: 4,57e-233 Length: 800
Score: 235.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.91% Indels: 0
DB: 4 Gaps: 0
US-10-751-736-84 (1-907) x BG196708 (1-800)
QY 404 AnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThrLeuProSer 423
DB 1 AATTTGGCTTGAACAAAATTGCTATTATTCACCCCAATGCATTTCCACITTTGCCATCC 60
QY 424 LeuIleLysLeuAspLysSerSerAsnLeuLeuSerSerPheProIleThrGlyLeuHis 443
DB 61 CTAATAAGCTGGACCTATCGTCCAACTTCCTGCTCTTTTCTTAACTAATCGGTACAT 120
QY 444 GlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerIleSerSerGlu 463
DB 121 GGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATATCATCTGAA 180
QY 464 AsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGly 483
DB 181 AACTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACAGTCTGTGCATTTGA 240
QY 484 ValCysGluAsnAlaTyrLysIleSerAsnGlnTTPAsnLysGlyAspAsnSerSerMet 503
DB 241 GTGTGTGAGAAATGCCTATAAGATTCTTAATCAATGGAATAAGGTGACAAACAGCATATG 300

```

```

QY 504 AspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGlu 523
DB 301 GACGACCTTCATAGAAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGTGACCTTGAA 360
QY 524 AspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerPro 543
DB 361 GATTTCTGCTTGACTTTGAGGAGACCTGAAAGCCCTTCATTCACTGACAGTGTTCACCT 420
QY 544 SerProGlyProPheLysProCysGluHisIleLeuLeuAspGlyTrpLeuIleArgIleGly 563
DB 421 TCCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGCTGTGATGGCTGATCAGAAATTGA 480
QY 564 ValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPhe 583
DB 481 GTGTGGACCATAGCAGTCTGCGACTTACTTGTAAATGCTTTGGTGACTTCAACAGTTTC 540
QY 584 ArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsn 603
DB 541 AGATCCCCCTCTGTACATTTCCCCCATTAACCTGTAATTGGGTTCATCGCAGCAGTGAAC 600
QY 604 MetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySer 623
DB 601 ATGCTCACGGGAGTCTCCAGTCCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
QY 624 PheAlaArgHisGlyAlaTyrTrpGluAsnGlyValGlyCysHis 638
DB 661 TTTGCACGACATGGTGGCTGGTGGGAGAAATGGGGTTGGTTGGCCAT 705
RESULT 4
LOCUS CD617354 590 bp mRNA linear EST 12-JAN-2004
DEFINITION S5041415H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD617354
VERSION CD617354.1 GI:40265619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.I.
AUTHORS Circular rapid amplification of cDNA ends for high-throughput
TITLE extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
Location/Qualifiers
source
1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.: 6,61e-175 Length: 590
Score: 179.00 Matches: 179
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.74% Indels: 0
DB: 6 Gaps: 0
US-10-751-736-84 (1-907) x CD617354 (1-590)
QY 98 AlaGlyAsnAlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLys 117
DB 16 GCGGAAACGCTCTGACATACATTCACCAAGGAGCAATTCACCTGGCGCTTTACAGTCTTAA 75

```


QY 118 ValLeuMetLeuGlnAnshnGlnLeuArgHisValProThrGluAlaLeuGlnAnshnLeu 137
 DB 76 GTTCTTATGCTGCAGATAATACGCTAAGACACGTAACCCACAGAGCTCTGCAGATTGG 135
 QY 138 ArgSerLeuGlnSerLeuArgLeuAspAlaAnshnHisLeuSerTyrValProProSerCys 157
 DB 136 CGAGGCTTCATCTCCCTGGTCTGGATGTATACCATCAGCTATGTGCCCCAGCTGT 195
 QY 158 PheSerGlyLeuHisSerLeuArgHisLeuTyrLeuAspAspAnshnAlaLeuThrGluLe 177
 DB 196 TTCAGTGGCTGCATCTCCCTGAGGCACCTGTGGTGATGACATCGTTAAACAGAAATC 255
 QY 178 ProValGlnAlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAnshnLys 197
 DB 256 CCGGTCAGGCTTTTAGAAGTTTATCGGCAATTCGCAAGCCATGACCTTGGCCCTGGAACAA 315
 QY 198 IleHisHisIleProAspTyrAlaPheGlyAnshnLeuSerSerLeuValValLeuHisLeu 217
 DB 316 ATACACCATACCATGAGCTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTCTACATCTC 375
 QY 218 HisAnshnArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGlu 237
 DB 376 CATAACAATAGAAATCCCTCGGGAAGAAATGCTTTGATGGGCTCCACAGCCCTAGAG 435
 QY 238 ThrLeuAspLeuAnshnTyrAnshnLeuAspGluPheProThrAlaIleArgThrLeuSer 257
 DB 436 ACTTTAGATTAAATPACAAATACCTTTGATGAATTCGCCACTGCAATTAGGACACTCTCC 495
 QY 258 AnshnLeuGluLeuGlyPheHisSerAnshnIleArgSerIleProGluLysAla 276
 DB 496 AACCTTAAAGAACTAGGATTTATAGCAACATATCAGGTCGATACCTGAGAAGCA 552

RESULT 5
 BX403219
 LOCUS
 DEFINITION
 BX403219 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 Clone CS0D1029YG09 5-PRIME, mRNA sequence.

ACCESSION
 BX403219
 VERSION
 BX403219.2 GI:46875014
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1. (bases 1 to 873)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30619059.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 10101.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?8=CS2AAW3ZE04_AW92A03_1&c=10101.f.

FEATURES
 Location/Qualifiers
 1. .873
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1029YG09"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

source

ORIGIN

Alignment Scores:
 Pred. No.: 1,3e-171 Length: 873
 Score: 176.00 Matches: 176
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 19.40% Indels: 0
 DB: 5 Gaps: 0

US-10-751-736-84 (1-907) x BX403219 (1-873)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
 DB 138 ATGGACACCTCCCGGCTCGGTGTGCTCTGCTCTTGCCTGTGCTGTGCTGAGCTGGCGACC 197
 QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
 DB 198 GGGGGCAGCTCTCCAGGCTGTGTGCTGAGGGGCTGCCCAACACACTGTCTATTGC 257
 QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
 DB 258 GAGCCCGACGGCAGGATGTTGCTCAGGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 317
 QY 61 ProSerAnshnLeuSerValPheThrSerTyrLeuAspLeuSerMetAnshnIleSerGln 80
 DB 318 COTTCCAACTCAGGCTCTTCACTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 377
 QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
 DB 378 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTGGAGGAGTTACGTTTTCGGGAAAC 437
 QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
 DB 438 GCTCTGACATACATTTCCCAAGGAGGACATTCCTCTGAGCTTTACAGTCTTAAGATTCATTG 497
 QY 121 LeuGlnAnshnGlnLeuArgHisValProThrGluAlaLeuGlnAnshnLeuArgSerLeu 140
 DB 498 CTGCAGAATAATCAGCTAAGACACGTCACCCACAGAAGCTCTGCAGAATTTGGAGAGCCTT 557
 QY 141 GlnSerLeuArgLeuAspAlaAnshnIleSerTyrValProProSerCysPheSerGly 160
 DB 558 CAATCCCTGCGTCTGGATGCTAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 617
 QY 161 LeuHisSerLeuArgHisLeuTyrLeuAspAspAnshnAlaLeuThrGlu 176
 DB 618 CTGCATTCTCCAGGACCTCTGGGCTGGATGACATGCGTTAACAGAA 665

RESULT 6

BX107244
 LOCUS
 DEFINITION
 BX107244 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
 IMAGE:795524, mRNA sequence.

ACCESSION
 BX107244
 VERSION
 BX107244.1 GI:27847308
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 475)
 Ebert, R., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
 Radelof, U., Schneider, D. and Korn, B.
 Human Unigeneset - RZPD3
 Unpublished (2003)
 Contact: Ina Rolfs

JOURNAL
 COMMENT
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGE:795524
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No.972)
 http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel.: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAACACGATGAC.

FEATURES

Location/Qualifiers
1..475

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998N171964 ; IMAGE:796624"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8 9w"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCGAAGCGGAGCGCGCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 3,85e-153 Length: 475
Score: 158.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.42% Indels: 0
DB: 5 Gaps: 0

US-10-751-736-84 (1-907) x BX107244 (1-475)

QY 566 ThrLeuAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585
DB 2 ACCATAGCAGTTCTGGCATTACTGTAATGCTTTGGTGACTTCAACAGTTTTCAGATCC 61
QY 586 ProLeuTyrLeuSerProLeuLeuLeuLeuLeuGlyValIleAlaAlaValAsnMetLeu 605
DB 62 CCTCTGTACATTTCCCCCATTAATCTGTTAATTTGGGGTCATCGCAGCAGTGAACATGCTC 121
QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625
DB 122 ACGGAGTCTCCAGTCCGCTGCTGGTGGTGGATGGTTCACCTTTTGGCAGCTTTCGA 181
QY 626 ArgHisGlyAlaTrrPrluAsnGlyValGlyCysHisValIleGlyPheLeuSerIle 645
DB 182 CGACATGGTCCCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCTATGGTTTTTTTGGCCATT 241
QY 646 PheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSer 665
DB 242 TTTGCTTCAGAAATCATCTGTTTCCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCCT 301
QY 666 ValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleIle 685
DB 302 GTGAATATTTCTGCAAAATTTGAACAGAAAGCTCCATTTCTAGCTGAAAGTATCATTT 361
QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLys 705
DB 362 TTGCTCTGTGCCCTGTGGCCCTTGACCATGGCGCAGTTCCCTCTGGGTGGCAGCAAG 421
QY 706 TyrGlyAlaSerProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 723
DB 422 TATGGGCCCTCCCTCTCTGCTGCTTTCCTTTTGGGGAGCCGCCAGCAATG 475

RESULT 7
AL530798

LOCUS

AL530798 1104 bp mRNA linear EST 24-MAR-2004
DEFINITION AL530798 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD008YF22 5-PRIME, mRNA sequence.

ACCESSION AL530798 GI:45705748

VERSION AL530798.3

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1104)

AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 13, 2001 this sequence version replaced gi:31068631.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7021.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DD008DC11QPI&c=7021.f.

FEATURES

source

Location/Qualifiers

1..1104

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DD008YF22"

/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 9,85e-152 Length: 1104
Score: 157.00 Matches: 312
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 2
Query Match: 17.31% Indels: 5
DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x AL530798 (1-1104)

QY 334 ThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnGlnLeuProAsn 353
DB 2 ACTTTAACTGGAGCAGACAGATCTCATCTCTCTCAACCCGTCTGCAATCAGTTACTTAAT 61
QY 354 LeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPheSerValCys 373
DB 62 CTCGAAGTGTAGATCTGCTTACAACTATTAGAAATTTACCCAGTTTTCAGTCTGC 121
QY 374 GlnLysLeuGlnLysIleAspLeuArgHisAsnGluIleTyrGluIleLysValAspThr 393
DB 122 CAAAAGCTTCAGAAAAATTGACCTTAAGACATAATGAATCTACGAATTTAAAGTTGCACACT 181
QY 394 PheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaTrrPheLysIleAlaIleIle 413
DB 182 TTCCAGCAGTTGCTTAGCTCCGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATT 241
QY 414 HisProAsnAlaPheSerThrLeuProSerLeuLysLeuAspLeuSerSerAsnLeu 433
DB 242 CACCCCAATGCAATTTTCCACTTTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTC 301
QY 434 LeuSerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsn 453

Db 302 CTGCTGCTCTTTCCTATTAACCTGGTTACATGGTTTAACTCACTTAAATTAACAGGAAT 361
 QY 454 HisAlaLeuGlnSerLeuIleSerSerGluAenPheProGluLeuIleValleGluMet 473
 Db 362 CATGCCCTTACAGAGCTTGATATCATCTGAAAACCTTCCAGAACTCAAGGTTATAGAAATG 421
 QY 474 Pro-TyrAlaTyrGlnCysCysAlaPheGlyValCysGlu-AenAlaTyrLys-IleSer 492
 Db 422 CCTATATGCTTACCAAGTGTGCTGTCATTTGGAGTGTGTTA-GAATGCTTATWA-GATTTCT 479
 QY 493 AenGlnTrpAsnLysGlyAspAenSerSerMetAspAspLeuHisLysLysAspAlaGly 512
 Db 480 NATCAATGGAATTAAGTTCACACAGCAGTATGGACGACTTCATTAAGAAAGATGCTGA 539
 QY 513 MetPheGlnAlaGlnAspGluArgAspLeuGluAaspPheLeuLeuAspPheGluLeuAep 532
 Db 540 ATGTTTTCAGGCTCAAGATGAAGCTGACCTTGAAGATTTCTGCTGACTTTGAGGAGAAC 599
 QY 533 LeuLysAlaLeuHisSerValGlnCysSerProSerProGlyProPheLysProCysGlu 552
 Db 600 CTGAAGCCCTTCATTCAGTCAGTGTGTTCCACTTCCCCAGGCCCTTCAACCCCTGTGAA 659
 QY 553 HisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThrIleAlaValLeuAlaLeu 572
 Db 660 CACCTGCTTGATGGCTGGCTGATCAGAAATGGAGTGTGACCATAGCAGTTCGGCACTT 719
 QY 573 ThrCysAsnAlaLeuValThrSerThrValPheArgSerProLeuTyrIleSerProIle 592
 Db 720 ACTTGTATGCTTTGGTGACTTCAACAGTCTTTCAGATCCCTCTGTATACATTTCCCCCAT 779
 QY 593 LysLeuLeuIleGlyValIleAlaValAenMetLeuThrGlyValSerSerAlaVal 612
 Db 780 AAACGTGTTAAATTTGGGGTTCATCCAGCAGTGAACATGCTCACGGAGTGTCCAGTCCCGTG 839
 QY 613 LeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyValaTrpTrpGlu 632
 Db 840 CTGGCTGGTGGTGGTGGCTTCACTTTTGGCAGCTTTGACGACATGTTGGCTGGTGGGAG 899
 QY 633 AenGlyValGlyCysHisValleGlyPheLeuSerIlePheAla 647
 Db 900 AATGGGTTGGTGGCATGTCTATGGTTTTTTTGTCCATTTTGGCC 944

RESULT 8
 AA460529 472 bp mRNA linear EST 09-JUN-1997
 LOCUS zxf8c09.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
 DEFINITION IMAGE:796624 5' similar to WP:C50H2.1 CE05479 STEROID HORMONE
 RECEPTOR ;, mRNA sequence.

ACCESSION AA460529 GI:2185649
 VERSION AA460529.1
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 472)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
 Moore,B., Schellenberg,K., Steptoe,M., Tan,P., Theising,B.,
 White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE WashU-Merck EST Project 1997
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LILNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
 source

Location/Qualifiers
 1. .472

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:6040989"
 /db_xref="taxon:9606"
 /clone="IMAGE:796624"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(GT) primer [5'
 TGTACCAATCTGAGTGGGAGCGCCGCTTAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Alignment Scores:
 Pred. No.: 4,62e-151 Length: 472
 Score: 156.00 Matches: 156
 Percent Similarity: 100.00% Conservatative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 17.20% Indels: 0
 DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x AA460529 (1-472)

QY 566 ThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585
 Db 3 ACCATAGCAGTCTCGGCACCTTACTTGTATATGCTTTGGTCACTCAACAGTTTTTCAGATCC 62
 QY 586 ProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAenMetLeu 605
 Db 63 CTCCTGTACATTTCCCCCATTAACCTGTAACTTTAATTTGGGGTCATCGCAGCAGTGAACATGCTC 122
 QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625
 Db 123 ACGGGAGTCTCCAGTCCGCTGCTGGCTGTGGATGCTTTCACATTTTGGCAGCTTTGCA 182
 QY 626 ArgHisGlyAlaTrpTrpGluAenGlyValGlyCysHisValleGlyPheLeuSerIle 645
 Db 183 CGACATGTGCTGTGTGGGAGAAATGGGGTGTGGTTCGCCATGTCAATTTGTTGCCATT 242
 QY 646 PheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSer 665
 Db 243 TTTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCT 302
 QY 666 VallysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValleIle 685
 Db 303 GTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCTCGAAGTATCATTT 362
 QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLys 705
 Db 363 TTGCTCTGTGCCCTGCTGCCCTTGACCATGGCCGAGTTCCTCCCTGCTGGTGGCAGCAG 422
 QY 706 TyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSer 721
 Db 423 TATGGCGCTCCCCCTCTCTGCTGCTTGGCTTTTGGGGAGCCCGCAGC 470

RESULT.9

LOCUS BQ323949 556 bp mRNA linear EST 17-MAY-2002
 DEFINITION IL5-CT0149-011100-224-a05 C10149 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BQ323949
 VERSION BQ323949.1 GI:20935938
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 556)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-CI0149-011100-224-a05&t3=2000-11-01&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 48.

FEATURES
 source
 1..556
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CI0149"
 /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 5,456-151 Length: 556
 Score: 156.00 Matches: 156
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 17.20% Indels: 0
 DB: 5 Gaps: 0

US-10-751-736-84 (1-907) x BQ323949 (1-556)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuGlnLeuAlaThr 20
 DB 88 ATGGACACCTCCGGCTCGGTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCGGACC 147
 QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
 DB 148 GGGGGGCGAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCAATTGC 207
 QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
 DB 208 GAGCCCGAGCGGAGGAGTGTCTCAGGGTGGACTGCTCCGACTCGGGGCTCTCGGAGCTG 267
 QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
 DB 268 CCTTCCACCTCAGCGCTCTTCCCTCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 327
 QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
 DB 328 CTGCTCCCGAATCCCTCCGCCAGTCTCCGCTTCTGGAGGAGGTAGCTCTTGGGGAAAC 387
 QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120

Db 388 GCTGTGCATACATTCCAGGGAGCATTCACCTGGCCTTTACAGTCTTAAAGTTCTTATG 447
 QY 121 LeuGlnAsnAsnGlnLeuArgGHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
 DB 448 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTCCGAAGCCTT 507
 QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSer 156
 DB 508 CAATCCCTGGCTCTGGATGCTTAACACATCAGCTATGTGCCCCAGC 555

RESULT 10
 BM795259
 LOCUS K-EST0077025 S21SNU520 Homo sapiens cDNA clone S21SNU520-31-C01 5',
 DEFINITION mRNA sequence.
 ACCESSION BM795259
 VERSION BM795259
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 671)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 31 row: C column: 01
 High quality sequence stop: 671.
 Location/Qualifiers
 1..671
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520-31-C01"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SNU520"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,278-143 Length: 671
 Score: 149.00 Matches: 222
 Percent Similarity: 99.11% Conservative: 0
 Best Local Similarity: 99.11% Mismatches: 1

REFERENCE 1 (bases 1 to 374)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1807 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 307.
FEATURES Location/Qualifiers
source
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:198741"
/tissue type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP_Utl"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

ORIGIN

Alignment Scores:
Pred. No.: 9,026-116 Length: 374
Score: 122.00 Matches: 122
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.45% Indels: 0
DB: 1 Gaps: 0
US-10-751-736-84 (1-907) x AI367723 (1-374)
Qy 675 LysAlaProPheSerLeuLysValIleLeuLeuCysAlaLeuAlaLeuThr 694
Db 372 AAAGTCCATTTCTAGCGCTGAAGTAATCATTTTGTCTGTGGCCCTGTGGCCTTGACC 313
Qy 695 MetAlaValProLeuLeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuPro 714
Db 312 ATGGCCGAGTTCCCTGCTGGTGGGAGCAAGTATGGGCTCCCTCTCTGCTGCT 253
Qy 715 LeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSer 734
Db 252 TTGCCTTTTGGGAGCCAGCACCATGGCTACATGGTCTCATCTTGTCTCAATTC 193
Qy 735 LeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGly 754
Db 192 CTTTGCTTCCCATGATGACCATTCCTTACCAAGCTCTACTGCAATTTGGACAAAGGA 133
Qy 755 AspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAlaLeuLeuPheThr 774
Db 132 GACCTGGAGATATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTTGTCTTAC 73
Qy 775 AsnCysIleLeuAsnCysProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThr 794
Db 72 AACTGCATCCTAAATGCGCCCTGTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCT 13
Qy 795 PheIle 796
Db 12 TTTATC 7

RESULT 13

AI699236/c
LOCUS AI699236 358 bp mRNA linear EST 16-DEC-1999
DEFINITION tx22f10.x1 NCI CGAP Utl1 Homo sapiens cDNA clone IMAGE:2274187 3'
similar to TR:O75473 O75473 ORPHAN G PROTEIN-COUPLED RECEPTOR HG38.
i, mRNA sequence.
ACCESSION AI699236
VERSION AI699236
KEYWORDS EST.
SOURCE AI699236.1 GI:4987136
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 3883 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 329.
FEATURES Location/Qualifiers
source
1..358
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2274187"
/tissue type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP_Utl"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
ORIGIN
Alignment Scores:
Pred. No.: 1,26e-111 Length: 358
Score: 118.00 Matches: 118
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.01% Indels: 0
DB: 1 Gaps: 0
US-10-751-736-84 (1-907) x AI699236 (1-358)
Qy 692 LysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeuLeu 701
Db 358 AAAGTAATCATTTTGTCTGTGGCCCTGTGGCCTTGACCATGGCGCAGTTCCCTGCTG 299
Qy 702 GlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSer 721
Db 298 GGTGGCAGCAAGTATGGGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCCAGC 239
Qy 722 ThrMetGlyTyrMetValAlaLeuLeuLeuLeuAsnSerLeuCysPheLeuMetThr 741
Db 238 ACCATGGGTCATAGTGGCTCTCATCTTGCTCAATTCCTTTGCTCTCATGTATGACC 179
Qy 742 IleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAsp 761
Db 178 ATTGCTTACCAAGCTCTACTGCAATTTGACAAAGGAGACCTGGAGATATTTGGGAC 119
Qy 762 CysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysPro 781


```

Db      118 TGGCTTATGGTAAACACATTGCCCTGTGGCTTTCACCAACTGCATCTAAACGCGCT 59
Qy      782 ValAlaPheLeuSerPheSerLeuLeuAenLeuThrPheIleSerProGlu 799
Db      58 GTGGCTTCTTCTGCTCTCTCTCTTATAAACCCTTACATTATCAGTCTGAA 5

RESULT 14
A1697103/c 539 bp mRNA linear EST 14-DEC-1999
LOCUS tql4f08.x1 NCI CGAP Ut3 Homo sapiens cDNA clone IMAGE:2208807 3'
DEFINITION similar to TR:075473 075473 ORPHAN G PROTEIN-COUPLED RECEPTOR HG38.
; mRNA sequence.
ACCESSION A1697103
VERSION A1697103.1 GI:4985003
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 960 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 405.
FEATURES
Location/Qualifiers
source
1..539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2208807"
/tissue_type="poorly-differentiated endometrial
adenocarcinoma, 2 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Ut3"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.45 kb. Life Technologies catalog #:
11541-018"

ORIGIN
Alignment Scores:
Pred. No.: 3..676-104 Length: 539
Score: 111.00 Matches: 111
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.24% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x A1697103 (1-539)

Qy      427 LeuAspLeuSerSerAsnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr 446
Db      334 CTGGACCTATCGTCCCAACTCTCGTCTCTTTCATTAACCTGGGTATCATGTTAACT 275
Qy      447 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 466
Db      274 CACTTAAATTAACAGGAATCATCGCTTACAGACTTGATATCATCTGAAACTTCCA 215
Qy      467 GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 486

```

```

Db      214 GAATCAAGGTATAGAAATGCTTATGCTTACCAGTGTGTCATTGGAGTGTGTGAG 155
Qy      487 AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 506
Db      154 AATGCTATATAGATTTCTTAATCAATGGATTAAGGTGACACAGCATGATGACGACCTT 95
Qy      507 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluAragAspLeuGluAspPheLeu 526
Db      94 CATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCCTG 35
Qy      527 LeuAspPheGluGluAspLeuLysAlaLeuHis 537
Db      34 CTTGACTTTGAGGAAGACCTGAAAGCCCTTCAT 2

RESULT 15
A1697103 2593 bp DNA linear GSS 16-DEC-2003
LOCUS Pan troglodytes GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY411733
VERSION AY411733.1 GI:39767701
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 2593)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2593)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..2593
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2593
/gene="GPR49"
/locus_tag="HCM4322"

ORIGIN
Alignment Scores:
Pred. No.: 6..03e-90 Length: 2593
Score: 98.00 Matches: 223
Percent Similarity: 98.24% Conservative: 0
Best Local Similarity: 98.24% Mismatches: 2
Query Match: 10.80% Indels: 4
DB: 9 Gaps: 0

US-10-751-736-84 (1-907) x AY411733 (1-2593)

Qy      610 SerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyAla 629
Db      1697 AGTGGCGTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1756
Qy      630 TrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGlu 649
Db      1757 TGGTGGGAGAAATGGGGTTGGTTGGCATGTCATGTTTTTGTCCATTTTGTCTTCGGA 1816

```

```
QY      650 SerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerValIysTyrSer 669
Db      1817 TCAICTGTTTTCTGCTTACTCTGGACGCCCTGGAGCGTGGGTCTCTGTGAATATTTCT 1876

QY      670 AlalysPheGluThrLysAlaProPheSerSerLeuLysValIleIleLeuLeuCysAla 689
Db      1877 GCAAAATTTGAAACGAAAGCTCCCTTTCTAGCCGTGAAAGTAATCATTTTGTCTGTGCC 1936

QY      690 LeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyrGlyAla-Se 709
Db      1937 CTGCTGGGCTTTGACCATGGCGCAGTTCCCTGCTGGGTGGCAGCAAGTATGGCAC-CTC 1995

QY      709 rProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLe 729
Db      1996 CCTCTCTGCGCTGCTTTGGCTTTTGGGGAGCCCGACCATGGGCTACATGGTGGCTCT 2055

QY      729 uIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCy 749
Db      2056 CATCTTGCTCAATTCCTTTGCTTCTCATGATGACCATTGCTACACCAAGCTCTACTG 2115

QY      749 aAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAl 769
Db      2116 CAATTGGACAAGGGAGACCTGGAGATATTGGGACTGCTCTATGGTAAAAACACATTGC 2175

QY      769 aLeuLeuLeuPheThrAsnCysIleLeuAsn-CysProValAlaPheLeuSerPheSerS 789
Db      2176 CCTGTTGCTCTTCACCAACTGCATCCTATA-CTGCCCTGTGGCTTCTTGTCTCTCCT 2234

QY      789 erLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeuValValV 809
Db      2235 CTTTAATAAACCTTACATTTTATCAGTCTGGAAGTAATTAAGTTTATCCTTCTGGTGGTAG 2294

QY      809 alProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHisPheLysG 829
Db      2295 TCCCACTTCTCGCATGTCTCAATCCCTTCTCTACATCCTGTTCATCCTCCTCACTTTAAGG 2354

QY      829 luAspLeuValSerLeu 834
Db      2355 AGGATCTGTGTAGGCTG 2371
```

Search completed: July 12, 2005, 12:55:36
Job time : 7011 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: July 12, 2005, 07:36:43 ; Search time 103 Seconds
(without alignments)
3405.744 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLGLVLSIPVLQLAT.....PAYPTESCHLSVAFVPCFL 907

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 2105692 seqs, 386760381 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : A_Geneseq_16Dec04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907	100.0	907	2	AAW93889 Human HG3
2	907	100.0	907	3	AAY90682 Human G p
3	907	100.0	907	6	ABP81968 Human G p
4	907	100.0	907	6	ABO06467 Human G-p
5	907	100.0	907	7	ADC22783 Human G p
6	907	100.0	907	7	AD559150 Human Pro
7	907	100.0	907	7	AD559153 Human Pro
8	907	100.0	907	7	ADG42628 Human G p
9	907	100.0	907	7	ADH14256 Human HG3
10	907	100.0	907	7	ADN40013 Cancer/an
11	907	100.0	907	7	ADN39531 Cancer/an
12	907	100.0	907	7	ADN39628 Cancer/an
13	907	100.0	907	8	ADO29408 Human GPC
14	907	100.0	907	8	ADG80369 G protein
15	907	100.0	907	8	ADR67868 Human HG3
16	806	88.9	907	3	AAY90687 Human mut
17	806	88.9	907	7	ADC22797 Human G p
18	806	88.9	907	7	ADH14270 Mutated h
19	806	88.9	1145	7	ADF70480 Orphan re
20	716	78.9	907	7	ADG42629 Human G p
21	689	76.0	692	2	AAW93890 Human HG3
22	621	68.5	883	7	ADB80464 Ovarian c
23	621	68.5	883	7	ADN40012 Cancer/an
24	621	68.5	883	7	ADN39166 Cancer/an
25	592	65.3	693	2	AAY42169 Human LGR

26	202	22.3	202	2	AAV53573	Human gon
27	197	21.7	282	5	ABP42000	Human ova
28	51	5.6	907	8	ADG75449	Mouse orp
29	51	5.6	907	8	ADO29409	Mouse GPC
30	43	4.7	43	2	AAW93892	Human HG3
31	43	4.7	43	2	AAW93894	Human HG3
32	43	4.7	43	2	AAW93893	Human HG3
33	43	4.7	43	2	AAW93895	Human HG3
34	43	4.7	43	2	AAW93896	Human HG3
35	43	4.7	43	2	AAW93891	Human HG3
36	20	2.2	20	6	ABP82972	G protein
37	20	2.2	20	6	ABP82976	G protein
38	18	2.0	91	2	AAV53577	Human gon
39	18	2.0	244	8	ADG66490	Novel hum
40	18	2.0	266	3	AAW41583	Human ORP
41	18	2.0	736	5	AAW66140	Human LGR
42	18	2.0	736	7	ADN02246	Human par
43	18	2.0	736	8	ADK19411	Human LGR
44	18	2.0	794	5	ADH50810	Human G-p
45	18	2.0	823	5	ADH50809	Human G-p

ALIGNMENTS

RESULT 1
AAW93889
ID AAW93889 standard; protein; 907 AA.

AC AAW93889;

DT 25-JUN-1999 (first entry)

DE Human HG38 protein.

KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
KW endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator.

OS Homo-sapiens.

PN WO9915660-A1-

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US019979.

XX 24-SEP-1997; 97US-0059863P.

XX (MERI) MERCK & CO INC.

XX Liu Q, Bailey WJ, McDonald TP;

XX WPI; 1999-254711/21.

XX N-PSDB; AAX23980.

XX Human G-protein coupled glycoprotein hormone receptor HG38.

XX Claim 1a; Fig 2; 74pp; English.

XX This invention describes a novel human G-protein coupled glycoprotein hormone receptor, HG38. Glycoprotein hormone receptors are important in the endocrine system and HG38 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the HG38 receptor, as well as for studying the ability of a variety of compounds to act as modulators of HG38 receptor activity

XX Sequence 907 AA;

Query Match 100.0%; Score 907; DB 2; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;

		Matches	907;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MDTSLGLVLSLPVLLQLATGSSPRSGVILRGCPHCHCEPDGMRLLRVDCSDLGSEL	60								
Ds	1	MDTSLGLVLSLPVLLQLATGSSPRSGVILRGCPHCHCEPDGMRLLRVDCSDLGSEL	60								
Qy	61	PSNLSVFTSYDLSMNNISQLLPNPLSLRFLFEEELRLAGNALTYIPKGAFTGLSKLVLM	120								
Ds	61	PSNLSVFTSYDLSMNNISQLLPNPLSLRFLFEEELRLAGNALTYIPKGAFTGLSKLVLM	120								
Qy	121	LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAALTEIPVQ	180								
Ds	121	LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAALTEIPVQ	180								
Qy	181	AFRSLALQAMTLALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCKPGLHSLETLID	240								
Ds	181	AFRSLALQAMTLALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCKPGLHSLETLID	240								
Qy	241	LYNNLDEFPPTAIRTLNKLGHFSNNIRIPEKAFVGNPISLITIHFDYDNPIDFVGRSA	300								
Ds	241	LYNNLDEFPPTAIRTLNKLGHFSNNIRIPEKAFVGNPISLITIHFDYDNPIDFVGRSA	300								
Qy	301	FOHLPRLRTLTLNGASQITEPDDITGTANLESRLTGAQISSLPQTVCNQPNLQVLDLS	360								
Ds	301	FOHLPRLRTLTLNGASQITEPDDITGTANLESRLTGAQISSLPQTVCNQPNLQVLDLS	360								
Qy	361	YNLLEDLPSFVCQKLOKIDIRHNEIYEIKVDTFQQLLSLSLNNKIAIHHNAPST	420								
Ds	361	YNLLEDLPSFVCQKLOKIDIRHNEIYEIKVDTFQQLLSLSLNNKIAIHHNAPST	420								
Qy	421	LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENFPELKVEMPVAYQCC	480								
Ds	421	LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENFPELKVEMPVAYQCC	480								
Qy	481	AFVCENAYKISNQNKGDNSSMDLHKDAGMFQAQDERDLEPLDFEEDLKALHSVQ	540								
Ds	481	AFVCENAYKISNQNKGDNSSMDLHKDAGMFQAQDERDLEPLDFEEDLKALHSVQ	540								
Qy	541	CSPSPGPKCEHLLDGLHILBIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600								
Ds	541	CSPSPGPKCEHLLDGLHILBIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600								
Qy	601	AVNMLTGVSAAVLGADAFSGSFARHGWNGVGHVIGFLSTIFASESSVFLTLAAL	660								
Ds	601	AVNMLTGVSAAVLGADAFSGSFARHGWNGVGHVIGFLSTIFASESSVFLTLAAL	660								
Qy	661	ERGSVKYSAXFETKAPSSLKVIILLCALLALTAAVPLLGSKYKASPLCLPLPFGEP	720								
Ds	661	ERGSVKYSAXFETKAPSSLKVIILLCALLALTAAVPLLGSKYKASPLCLPLPFGEP	720								
Qy	721	STMGYVALIILNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780								
Ds	721	STMGYVALIILNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780								
Qy	781	PVAFLSFSSLLNLTFISPEVIKFIILVVVPLPACINPLLYILFNPFKEDLVSLRKOTYV	840								
Ds	781	PVAFLSFSSLLNLTFISPEVIKFIILVVVPLPACINPLLYILFNPFKEDLVSLRKOTYV	840								
Qy	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTTSSITVDLPSPSPVPAYPVTESCHLSS	900								
Ds	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTTSSITVDLPSPSPVPAYPVTESCHLSS	900								
Qy	901	VAFVPCVL 907									
Ds	901	VAFVPCVL 907									

RESULT 2
RAY90682
ID AAY90682 standard; protein; 907 AA.
XX
AC
XX

DT	21-AUG-2000	(first entry)
XX	Human G protein-coupled receptor HG38.	
DE	G protein-coupled receptor; GPCR; constitutively active;	
KW	intracellular loop 3; transmembrane domain 6; drug screening; agonist;	
KW	antagonist.	
XX	Homo sapiens.	
OS	WO200022129-A1.	
PN	20-APR-2000.	
PD	12-OCT-1999; 99WO-US023938.	
PF	13-OCT-1998; 98US-00170496.	
PR	(AREN-) ARENA PHARM INC.	
XX	Behan DP, Chalmers DT, Liaw CW;	
PA	WPI; 2000-329165/28.	
XX	N-PSDB; AAA30770.	
DR	Non-endogenous constitutively activated human G protein-coupled	
PT	receptors, useful for identifying agonists for use as pharmaceutical	
PT	agents.	
XX	Example 1; Page 317-320; 341pp; English.	
PS	The invention relates to constitutively active, non-endogenous versions	
XX	of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-	
CC	AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743	
CC	and AAA30775-A30779). The mutant proteins of the invention contain a	
CC	mutation in a portion of the protein comprising intracellular loop 3	
CC	(IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,	
CC	is substituted for an endogenous residue in IC3 at a position 16 amino	
CC	acids N-terminal of an endogenous proline in TM6 to form a sequence X-	
CC	(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or	
CC	Ala, and is preferably Lys. When the endogenous residue at this position	
CC	is Lys, this residue is replaced by His, Arg or preferably Ala. The 15	
CC	amino acid stretch between the substituted amino acid and the Pro may be	
CC	endogenous, non-endogenous, or a mixture of endogenous and non-endogenous	
CC	residues. The constitutively active GPCRs are useful for identifying	
CC	antagonists, agonists and partial agonists for use as pharmaceutical	
CC	agents. The mutant proteins are also useful in research settings for	
CC	elucidating the roles of the receptors in normal and diseased conditions.	
CC	Antagonists for a particular GPCR are useful for treating diseases and	
CC	disorders associated with that receptor. Because the novel mutant GPCRs	
CC	are constitutively active, they can be used directly for screening of	
CC	compounds without the need for endogenous ligands. The present sequence	
CC	represents a human wild-type GPCR referred to in an exemplification of	
CC	the invention	
XX	Sequence 907 AA:	
SO		

Qy	1	MDTSLGLVLSLPVLLQLATGSSPRSGVILRGCPHCHCEPDGMRLLRVDCSDLGSEL	60
Ds	1	MDTSLGLVLSLPVLLQLATGSSPRSGVILRGCPHCHCEPDGMRLLRVDCSDLGSEL	60
Qy	61	PSNLSVFTSYDLSMNNISQLLPNPLSLRFLFEEELRLAGNALTYIPKGAFTGLSKLVLM	120
Ds	61	PSNLSVFTSYDLSMNNISQLLPNPLSLRFLFEEELRLAGNALTYIPKGAFTGLSKLVLM	120
Qy	121	LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAALTEIPVQ	180
Ds	121	LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAALTEIPVQ	180

QY 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVLHLHNNRIHSLGKKCFDGLHSLETLD 240
DB 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVLHLHNNRIHSLGKKCFDGLHSLETLD 240
QY 241 LNYNLDLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300
DB 241 LNYNLDLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEPDLTGNTANLSLTLTGAISSLPQVCNQLPQLVLDLS 360
DB 301 FOHLPELRTLTLNGASQITTEPDLTGNTANLSLTLTGAISSLPQVCNQLPQLVLDLS 360
QY 361 YNLLEDLPSFVCQKLOKIDLRHNEIYKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPSFVCQKLOKIDLRHNEIYKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKDLNLSNLSRPFITGLHGTUHLKTCGNHALQSLISSENPELKIEMPVAYOCC 480
DB 421 LPSLIKDLNLSNLSRPFITGLHGTUHLKTCGNHALQSLISSENPELKIEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPPKPEHLLDGLMIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
DB 541 CSPSPGPPKPEHLLDGLMIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVCVIGFLSTFASSESVELLTLAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVCVIGFLSTFASSESVELLTLAAL 660
QY 661 ERGSFVKYSAKFETKAPFSSLVKVIILLCALLALTAALMAAVPLLGSGKYGASPLCLPFGEP 720
DB 661 ERGSFVKYSAKFETKAPFSSLVKVIILLCALLALTAALMAAVPLLGSGKYGASPLCLPFGEP 720
QY 721 STMGVMVALIILNLSLCFLMNTIATKLYCNLDKGDLENWDCSMVKHIALLLFNCLNC 780
DB 721 STMGVMVALIILNLSLCFLMNTIATKLYCNLDKGDLENWDCSMVKHIALLLFNCLNC 780
QY 781 PVAFLSFSSNLNLFISPEVIFKILLVVVLPACLNPLLYLTFHPFKEDLVLSRKQYV 840
DB 781 PVAFLSFSSNLNLFISPEVIFKILLVVVLPACLNPLLYLTFHPFKEDLVLSRKQYV 840
QY 841 WTRSGHPSLMSINDVVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
DB 841 WTRSGHPSLMSINDVVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

RESULT 3

ABP81968
ID ABP81968 standard; protein; 907 AA.

AC ABP81968;

DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor GPR49 protein SEQ ID NO:422.

KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

ulcer.

OS Homo sapiens.

XX WO200261087-A2.

XX 08-AUG-2002.

XX 19-DEC-2001; 2001WO-US050107.

XX 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burmer GC, Roush CL, Brown JP;

XX WPI; 2003-046718/04.

XX N-PSDB; ABZ42816.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

PS Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Sequence 907 AA;

Query Match 100.0%; Score 907; DB 6; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGLVLLSLPVLLQLATGSSPRSGVTLRGCPTHCHCEPDGMLRVDGSDGLSEL 60
DB 1 MDTSLRGLVLLSLPVLLQLATGSSPRSGVTLRGCPTHCHCEPDGMLRVDGSDGLSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRPLEELRAGNALTYPKGAFTGLYSKVLV 120
DB 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRPLEELRAGNALTYPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSLSLQSLRLDANHSIYVPPSCFSGLSLRLHLDNALTETIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRSLSLQSLRLDANHSIYVPPSCFSGLSLRLHLDNALTETIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVLHLHNNRIHSLGKKCFDGLHSLETLD 240
DB 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVLHLHNNRIHSLGKKCFDGLHSLETLD 240

QY 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPPIQVGRSA 300
DB 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPPIQVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLOVLDS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLOVLDS 360
QY 361 YNLLEDLPFSVFCVKLOKIDLRHNEIYEIKVDYTFQQLLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVFCVKLOKIDLRHNEIYEIKVDYTFQQLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLSPPITGLHGTUHLKIGNHALQSLISSENFPELKVEMPVAYOCC 480
DB 421 LPSLIKLDLSSNLSPPITGLHGTUHLKIGNHALQSLISSENFPELKVEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKDDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPEHLLDGLHIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPKPEHLLDGLHIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPARHCAWENGVCVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSPARHCAWENGVCVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGSVKYSKAFETKAPFSSKLVIIICALLALTMAAVPLLGSGKYGASPLCLPFGEP 720
DB 661 ERGSVKYSKAFETKAPFSSKLVIIICALLALTMAAVPLLGSGKYGASPLCLPFGEP 720
QY 721 STMGVWALILNLSICFLMTIATYKLYCNLDKGBLENWDCSWKHYHALLFTNCILNC 780
DB 721 STMGVWALILNLSICFLMTIATYKLYCNLDKGBLENWDCSWKHYHALLFTNCILNC 780
QY 781 PVAFLSFSLNLTPISEVVKIFILLVVPACLNPLLYILFNPHKEDLVSLRKQTYV 840
DB 781 PVAFLSFSLNLTPISEVVKIFILLVVPACLNPLLYILFNPHKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSITYDLPPSSVPSPAYPVTTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTTSSITYDLPPSSVPSPAYPVTTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

RESULT 4

ABO06467

ID ABO06467 standard; protein; 907 AA.

XX

AC ABO06467;

XX

XX 13-AUG-2003 (first entry)

XX

DE Human G-protein coupled receptor HG38.

XX

KW G-protein coupled receptor; GPCR; caudate nucleus related disorder;
KW neoplastic disorder; leukaemia; breast cancer; immune disorder; AIDS;
KW rheumatoid arthritis; neuronal disorder; Alzheimer's disease; receptor;
KW Parkinson's disease; respiratory disorder; bronchopulmonary disease;
KW pancreatic disease; ovarian disorder; ovarian carcinoma; colon disease;
KW Meig's syndrome; breast disease; thalamus related disorder; gene therapy;
KW amygdala related disorder; corpus callosum related disorder;
KW hippocampus related disorder.

XX Homo sapiens.

OS

XX US2003027323-A1.

XX

PN

PD 06-FEB-2003.

XX 26-SEP-2001; 2001US-00965536.

XX 27-SEP-2000; 2000US-0235832P.

PR 16-JAN-2001; 2001US-0261781P.

PR 19-JUL-2001; 2001US-030605P.

PR 03-AUG-2001; 2001US-0310436P.

XX (PEDE/) FEDER J N.

PA (MINT/) MINTIER G.

PA (RAMA/) RAMANATHAN C S.

PA (HAWK/) HAWKEN D R.

XX Feder JN, Mintier G, Ramanathan CS, Hawken DR;

XX WPI; 2003-466147/44.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

h.s. e. 102e

Rab

102e

New G-protein coupled receptor polypeptides, designated as HGPRBMY5, useful for preventing, treating or ameliorating a medical condition related to the colon, breast, ovaries or immune system.

Disclosure; Fig 8; 93pp; English.

The invention relates to an isolated HGPRBMY5 polypeptide. The polypeptides, polynucleotides and methods are useful for preventing, treating or ameliorating a medical condition such as a neoplastic disorder e.g. leukaemia and breast cancer; immune disorder e.g. AIDS and rheumatoid arthritis; neuronal disorder e.g. Alzheimer's disease and Parkinson's disease; respiratory disorder e.g. bronchopulmonary disease and pancreatic's syndrome; ovarian disorder e.g. ovarian carcinoma and Meig's syndrome; colon disease; breast disease; thalamus related disorder; amygdala related disorder; corpus callosum related disorder; caudate nucleus related disorder; hippocampus related disorder by administering the GPCR polypeptide or its homologue. The present sequence represents the amino acid sequence of a G-protein coupled receptor used to show homology with the human G-protein coupled receptor, HGPRBMY5.

Sequence 907 AA;

Query Match 100.0%; Score 907; DB 6; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
DB 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRELELRAGNALTYIPKGAFTGLYSKVLV 120
DB 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRELELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSCPSGLHSRLHLDNALTETIYVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSCPSGLHSRLHLDNALTETIYVQ 180
QY 181 AFRSLSAQAMTALNKHITHIPDYAFGNLSLVVLHNNRHISLGKCKFGLHSLETLD 240
DB 181 AFRSLSAQAMTALNKHITHIPDYAFGNLSLVVLHNNRHISLGKCKFGLHSLETLD 240
QY 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPPIQVGRSA 300
DB 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPPIQVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLOVLDS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLOVLDS 360
QY 361 YNLLEDLPFSVFCVKLOKIDLRHNEIYEIKVDYTFQQLLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVFCVKLOKIDLRHNEIYEIKVDYTFQQLLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLSPPITGLHGTUHLKIGNHALQSLISSENFPELKVEMPVAYOCC 480

Db 421 LPSLKLQSSNLSFPITGLHGLTHLKTGNHALQSLSSSENPELKVEMPAYOCC 480
Qy 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERLDEFLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERLDEFLDPEEDLKALHSVQ 540
Qy 541 CSPSPGPPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLGADFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLGADFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAPKPTKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFPGBP 720
Db 661 ERGFSVKYSAPKPTKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFPGBP 720
Qy 721 STMGYVVALILNLSLFLMTIATYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALILNLSLFLMTIATYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFISPSLSINLTISPEVIFKILLVVPILPACLNPLLYILPNPKEDLVSLRKQTV 840
Db 781 PVAFISPSLSINLTISPEVIFKILLVVPILPACLNPLLYILPNPKEDLVSLRKQTV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPECL 907
Db 901 VAFVPECL 907
RESULT 5
ID ADC22783 standard; protein; 907 AA.
XX AC ADC22783;
XX DT 18-DEC-2003 (first entry)
XX DE Human G protein-coupled receptor (GPCR) polypeptide #39.
XX KW Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6;
XX intracellular-3 region; IC3; receptor..
XX OS Homo sapiens.
XX PN US6555339-B1.
XX PD 29-APR-2003.
XX PF 13-OCT-1998; 98US-00170496.
XX PR 14-APR-1997; 97US-00839449.
XX PR 14-APR-1998; 98US-00060188.
XX PR 26-JUN-1998; 98US-0090783P.
XX PR 07-AUG-1998; 98US-0095677P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-742861/70.
XX DR N-PSDB; ADC22782.
XX PT Creating a constitutively active version of an endogenous human G protein
XX coupled receptor (GPCR) comprises substituting a specific amino acid in
XX the transmembrane-6 region with a different amino acid, and testing for
XX constitutive activity.

XX Example 1; SEQ ID NO 264; 221pp; English.
XX The invention relates to a method for treating a non-endogenous,
CC constitutively active version of an endogenous human G protein-coupled
CC receptor (GPCR) that has a transmembrane-6 (TM6) region and an
CC intracellular-3 (IC3) region, by substituting a specific amino acid in
CC the TM6 region with a different amino acid, and testing for constitutive
CC activity. The method is useful for creating a constitutively active
CC version of an endogenous human GPCR that comprises a transmembrane 6
CC region and an intracellular loop 3 region. The altered human GPCR
CC polypeptides are useful for screening test compounds for identification
CC of inverse agonists or partial agonists of GPCR polypeptides, which may
CC have therapeutic uses. The altered GPCRs may also be used in vivo or in
CC vitro in biological research. A nucleic acid encoding the altered GPCR
CC may be used to create a transgenic animal expressing the altered GPCR.
CC The method allows screening for compounds that modulate the activity of a
CC human G protein-coupled receptor without the need for provision of a
CC ligand for the receptor. This is particularly useful in allowing
CC screening of compounds against orphan receptors for which no ligand is
CC currently known. This sequence represents a human GPCR polypeptide of the
CC invention.
XX Sequence 907 AA;
Qy Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDTSELGVLISLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRLRVDCSDGLSEL 60
Db 1 MDTSELGVLISLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRLRVDCSDGLSEL 60
Qy 61 PSNLSVFTSYLDLSNNISQLLPNPLPSIRFLFEEELRAGNALTYIPKGAFTGLYSIKVL 120
Db 61 PSNLSVFTSYLDLSNNISQLLPNPLPSIRFLFEEELRAGNALTYIPKGAFTGLYSIKVL 120
Qy 121 LQNLQHVPTALQNLRSLSQSLRDANHSYVPPSCFSGLSLHRLMDDNALTEIPVQ 180
Db 121 LQNLQHVPTALQNLRSLSQSLRDANHSYVPPSCFSGLSLHRLMDDNALTEIPVQ 180
Qy 181 AFRSLSALQMTALANKIHHIPDYAFGNLSLVHLHNNRHSIGKCKFDGLHSL 240
Db 181 AFRSLSALQMTALANKIHHIPDYAFGNLSLVHLHNNRHSIGKCKFDGLHSL 240
Qy 241 LNYNMLDEFPFAIRTLNKLKELGFHNNIRISPEKAFVGNPSLTIHFDNPIQFVGRSA 300
Db 241 LNYNMLDEFPFAIRTLNKLKELGFHNNIRISPEKAFVGNPSLTIHFDNPIQFVGRSA 300
Qy 301 FOHLPELRTLINGASQITEFPDLTGTLNLESLSLTGAQISSLPQVCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLINGASQITEFPDLTGTLNLESLSLTGAQISSLPQVCNQLPNLQVLDLS 360
Qy 361 YNLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFQQLLSLSLNLNLANWKIATIHNAFST 420
Db 361 YNLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFQQLLSLSLNLNLANWKIATIHNAFST 420
Qy 421 LPSLIKLDSNLLSSFPITGLHGLTHLKTGNHALQSLSSSENPELKVEMPAYOCC 480
Db 421 LPSLIKLDSNLLSSFPITGLHGLTHLKTGNHALQSLSSSENPELKVEMPAYOCC 480
Qy 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERLDEFLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERLDEFLDPEEDLKALHSVQ 540
Qy 541 CSPSPGPPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLGADFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLGADFTGSGFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660

Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
Db |||||
Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
Db |||||
Qy 721 STMGVMVALILLNSICFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db |||||
Qy 781 PVAFSLFSSLINLTPISEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db |||||
Qy 781 PVAFSLFSSLINLTPISEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTSSCHLSS 900
Db |||||
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTSSCHLSS 900
Qy 901 VAFVPCCL 907
Db |||||
Qy 901 VAFVPCCL 907

RESULT 6
ADE59150
ID ADE59150 standard; protein; 907 AA.

AC ADE59150;

DT 29-JAN-2004 (first entry)

DE Human Protein O75473, SRQ ID NO 5041.

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; O75473.

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLGVLLSLPVLLQLATGCGSSPRSGVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
Db |||||
Qy 1 MDTSRGLGVLLSLPVLLQLATGCGSSPRSGVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
Db |||||
Qy 61 PSNLSVFTSYLDLSNMNISOQLLPNPLPSLRFLBELRLAGNALTYIPKGAFTGLYSKVLV 120
Db |||||
Qy 61 PSNLSVFTSYLDLSNMNISOQLLPNPLPSLRFLBELRLAGNALTYIPKGAFTGLYSKVLV 120
Db |||||
Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHISYVPPSCFSGLSLRLHLDLNNALTEIPVQ 180
Db |||||
Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHISYVPPSCFSGLSLRLHLDLNNALTEIPVQ 180
Db |||||
Qy 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCKPGLHSLEFLD 240
Db |||||
Qy 181 AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCKPGLHSLEFLD 240
Db |||||
Qy 241 LNYNNLDEFFTAIRTLNKLGHPSNNIRISPEKAFVGNPSLTIHFYDNPIDQVGRSA 300
Db |||||
Qy 241 LNYNNLDEFFTAIRTLNKLGHPSNNIRISPEKAFVGNPSLTIHFYDNPIDQVGRSA 300
Db |||||
Qy 301 FOHLPELRTLNGASQITEPPDLTGANLESILTLTGAQISSLPQTVCNQLPQLVDLS 360
Db |||||
Qy 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLLSLRLANNKIAIHPNAPST 420
Db |||||
Qy 361 YNLLLEDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLLSLRLANNKIAIHPNAPST 420
Db |||||
Qy 421 LPSLIKLDLSSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIYVQCC 480
Db |||||
Qy 421 LPSLIKLDLSSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMPIYVQCC 480
Db |||||
Qy 481 AFGVCENAYKISNOWNKGDSSMDLHKDKAGMFQADDERDLEFLDDEFEDLKALHSVQ 540
Db |||||
Qy 481 AFGVCENAYKISNOWNKGDSSMDLHKDKAGMFQADDERDLEFLDDEFEDLKALHSVQ 540
Db |||||
Qy 541 CSPSPGPKCEHLLDGLWIRIGVWTTAVLTATCNALVTSTVFRSPLYSIPKLLIGVIA 600
Db |||||
Qy 541 CSPSPGPKCEHLLDGLWIRIGVWTTAVLTATCNALVTSTVFRSPLYSIPKLLIGVIA 600
Db |||||
Qy 601 AVNMLTGVSAAVLAVGDAFTFGSPARHGAWNGVGHVIGFLSIFASESSVFLTLAAL 660
Db |||||
Qy 601 AVNMLTGVSAAVLAVGDAFTFGSPARHGAWNGVGHVIGFLSIFASESSVFLTLAAL 660
Db |||||
Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
Db |||||
Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
Db |||||
Qy 721 STMGVMVALILLNSICFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db |||||
Qy 721 STMGVMVALILLNSICFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db |||||

QY 781 PVAFSLFSSLLINLTFTSPVKEIKLVVPLACINPLIYILFNPHFKEDLYSLRKQTVY 840
 DB 781 PVAFSLFSSLLINLTFTSPVKEIKLVVPLACINPLIYILFNPHFKEDLYSLRKQTVY 840
 QY 841 WTRSKHPSLMSINSDVEKQCDSTQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900
 DB 841 WTRSKHPSLMSINSDVEKQCDSTQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900

QY 901 VAFVPCFL 907
 DB 901 VAFVPCFL 907

RESULT 7

AD59153
 ID AD59153 standard; protein; 907 AA.

XX AC AD59153;

DT 29-JAN-2004 (first entry)

XX Human Protein O75473, SEQ ID NO 5044.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; Chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; O75473.

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic form directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLGVLVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLVDCSDGLSEL 60

DB 1 MDTSLGVLVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLVDCSDGLSEL 60

QY 61 PNLNVFTSYLDLSMNNISQLLPNPLPSLRFUEELRAGNALTYPKGAFDTGLYSUKVLM 120

DB 61 PNLNVFTSYLDLSMNNISQLLPNPLPSLRFUEELRAGNALTYPKGAFDTGLYSUKVLM 120

QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALEIPVQ 180

DB 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALEIPVQ 180

QY 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLETLD 240

DB 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLETLD 240

QY 241 LNNYNNLDEFPPTAIRTLNKLGEHNNIRSRPEKAFVGNPNSLTIHFVDNPIQFVGRSA 300

DB 241 LNNYNNLDEFPPTAIRTLNKLGEHNNIRSRPEKAFVGNPNSLTIHFVDNPIQFVGRSA 300

QY 301 FOHLPELRTLINGASQITEFPDLTGATANLESLLTGAQISSLPQTCNQLNQLVDLS 360

DB 301 FOHLPELRTLINGASQITEFPDLTGATANLESLLTGAQISSLPQTCNQLNQLVDLS 360

QY 361 YNLLSDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420

DB 361 YNLLSDLPFSVCQKQKIDLRHNEIYEIKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420

QY 421 LPSLIKLDSLNLSSFFITGLHGLTHLKLTCNHALQSLISSENPELKVIEPVAQCC 480

DB 421 LPSLIKLDSLNLSSFFITGLHGLTHLKLTCNHALQSLISSENPELKVIEPVAQCC 480

QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEDFLLDPEEDLKALHSVQ 540

DB 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEDFLLDPEEDLKALHSVQ 540

QY 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

DB 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

QY 601 AVNMLTGSSAVLAGVDAFTGSPFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660

DB 601 AVNMLTGSSAVLAGVDAFTGSPFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660

QY 661 ERGFSVKYSAPKFKAPETKAPSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPFGGP 720

DB 661 ERGFSVKYSAPKFKAPETKAPSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPFGGP 720

QY 721 STMGYMVALILLNSLCFLMWTIATKLYCNLDKGLDENIWCMSVKHIALLLFTNCILNC 780

DB 721 STMGYMVALILLNSLCFLMWTIATKLYCNLDKGLDENIWCMSVKHIALLLFTNCILNC 780

QY 781 PVAFSLFSSLLINLTFTSPVKEIKLVVPLACINPLIYILFNPHFKEDLYSLRKQTVY 840

DB 781 PVAFSLFSSLLINLTFTSPVKEIKLVVPLACINPLIYILFNPHFKEDLYSLRKQTVY 840

QY 841 WTRSKHPSLMSINSDVEKQCDSTQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900

DB 841 WTRSKHPSLMSINSDVEKQCDSTQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900

QY	901	VAFPVCL 907	Db	181	AFRSLSALQAMTLALNKIHIPDYAFGNLSLVLHLHNNRHSIGKGCDFGLHSLETLD	240
Db	901	VAFPVCL 907	QY	241	LYNNLDEFPPTAIRTLNLKELGFHSNNIRSIPEKAFVGNPSLITIHFDYDNPQVGRSA	300
			Db	241	LYNNLDEFPPTAIRTLNLKELGFHSNNIRSIPEKAFVGNPSLITIHFDYDNPQVGRSA	300
RESULT 8						
ID	ADG42628		QY	301	FOHLPRLTLTLNGASQITTEPPDLTGTAANLESRLTGAQISSLPQTVCNQLPNLQVLDLS	360
XX	ADG42628	standard; protein; 907 AA.	Db	301	FOHLPRLTLTLNGASQITTEPPDLTGTAANLESRLTGAQISSLPQTVCNQLPNLQVLDLS	360
AC	ADG42628;		QY	361	YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDVTFQQLLSLRSLNLANMKIAIHPNAPST	420
DT	26-FEB-2004	(first entry)	Db	361	YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDVTFQQLLSLRSLNLANMKIAIHPNAPST	420
XX	Human G protein-coupled receptor 49.		QY	421	LPSLIKLDLSNLSLSPFITGLHGLTHLKLTCGNHALQSLISSENFPELKVTEMPEYVOC	480
KW	cytostatic; gene therapy; NOVX-agonist; NOVX-antagonist; pharmaceutical; NOVX-associated disorder; cancer; human; G protein coupled receptor 49.		Db	421	LPSLIKLDLSNLSLSPFITGLHGLTHLKLTCGNHALQSLISSENFPELKVTEMPEYVOC	480
XX	Homo sapiens.		QY	481	AFGVCENAYKISNOWNKGDNSMDLHKKQAGMFQAQDERDLEDFLLDFEEDLKALHSVQ	540
PN	US2003204052-A1.		Db	481	AFGVCENAYKISNOWNKGDNSMDLHKKQAGMFQAQDERDLEDFLLDFEEDLKALHSVQ	540
XX	30-OCT-2003.		QY	541	CSPSPGPFKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
PD	04-OCT-2001; 2001US-00970944.		Db	541	CSPSPGPFKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
XX	04-OCT-2000; 2000US-0237862P.		QY	601	AVNMLTGVSAAVLAVDAFTFGSFAHGAHWENGCHVIGFLSTFASESVELLTAAAL	660
PA	(HERR/) HERRMANN J L.		Db	601	AVNMLTGVSAAVLAVDAFTFGSFAHGAHWENGCHVIGFLSTFASESVELLTAAAL	660
PA	(RAST/) RASTELLI L.		QY	661	ERGFVKYSKAFETKAPFSSLKVIILLCALALATMAAVPLLGSGSKYGASPLCLPLPGEP	720
PA	(SHIM/) SHIMKETS R A.		Db	661	ERGFVKYSKAFETKAPFSSLKVIILLCALALATMAAVPLLGSGSKYGASPLCLPLPGEP	720
PI	Herrmann JL, Rastelli L, Shinkets RA;		QY	721	STMGYVVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENIWDCSVWKHIALLLFTNCILNC	780
DR	WPI; 2003-900673/82.		Db	721	STMGYVVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENIWDCSVWKHIALLLFTNCILNC	780
XX	New NOVX gene or NOVX-specific antibody, useful for preparing a		QY	781	PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	840
PT	composition for treating or preventing a NOVX-associated disorder, e.g.,		Db	781	PVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	840
PT	cancer.					
XX	Disclosure; SEQ ID NO 26; 118pp; English.					
PS	The invention describes a new isolated polypeptide comprising: a					
CC	polypeptide or its mature form comprising a sequence not given in the					
CC	specification; or a variant of (A), where one or more amino acid residues					
CC	in the variant differs in no more than 15% from the amino acid sequence					
CC	of the mature form. The pharmaceutical composition may be administered					
CC	via oral, transdermal, rectal or parenteral route. The polypeptide,					
CC	nucleic acid or antibody is useful for preparing a composition for					
CC	treating or preventing a NOVX-associated disorder, e.g., cancer. This is					
CC	the amino acid sequence of a transmembrane receptor homologue used in a					
CC	comparison with the novel human proteins of the invention.					
XX	Sequence 907 AA;					
SQ	Query Match 100.0%; Score 907; DB 7; Length 907;					
	Best Local Similarity 100.0%; Pred. No. 0;					
	Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MDTSRLGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60			
Db	1	MDTSRLGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60			
QY	61	PSNLSVFTSYDLSNWNISQLLPNPLPSLFLFEEURLAGNALTYIPKGAFTGLYSKVLVM	120			
Db	61	PSNLSVFTSYDLSNWNISQLLPNPLPSLFLFEEURLAGNALTYIPKGAFTGLYSKVLVM	120			
QY	121	LQNNQLRHVPTEALQNLSRLDANHSIVVPSCFSGLSLRLWLDNALTEIPVQ	180			
Db	121	LQNNQLRHVPTEALQNLSRLDANHSIVVPSCFSGLSLRLWLDNALTEIPVQ	180			
QY	181	AFRSLSALQAMTLALNKIHIPDYAFGNLSLVLHLHNNRHSIGKGCDFGLHSLETLD	240			

PR 13-OCT-1998; 98US-00170496.
XX (LIAW/) LIAW C W.
PA (BEHA/) BEHAN D P.
PA (CHAL/) CHALMERS D T.
XX
XX Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-801247/75.
DR N-PSDB; ADH14255.
XX
XX New constitutively active, non-endogenous version of an endogenous human
PT G protein-coupled receptor for the identification of therapeutic
PT compounds, such as agonists.
XX
XX Example 1; SEQ ID NO 264; 227pp; English.
XX
XX The invention relates to a constitutively active, non-endogenous version
CC of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC used for screening therapeutic compounds as inverse agonists, agonists or
CC partial agonists. The GPCR can be also used to elucidate and
CC understand the roles of GPCRs in normal and diseased humans. The GPCR
CC need not be purified and isolated to be used to screen for therapeutic
CC compounds. The utility of the GPCR as a research tool is enhanced because
CC the role of a particular receptor can be understood before the endogenous
CC ligand is identified. The present sequence is used in the exemplification
CC of the present invention.
XX
XX Sequence 907 AA;
XX
Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRIGVLLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSL 60
DB 1 MDTSRIGVLLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSL 60
QY 61 PNLNSVFTSYLDLNMNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLNSVFTSYLDLNMNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
QY 121 LQNNQLRHVPTEALQNLRLQSLRLDANHSVPPSPCFSGLSLRLHLDLDDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRLQSLRLDANHSVPPSPCFSGLSLRLHLDLDDNALTEIPVQ 180
QY 181 AFRSLALQAMTLANKIHHI PDYAFGNLSLVVLHNNRTHSLGKKCFDGLHSL 240
DB 181 AFRSLALQAMTLANKIHHI PDYAFGNLSLVVLHNNRTHSLGKKCFDGLHSL 240
QY 241 LNNYNDLDEPTAIRTLNLSKELGFSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
DB 241 LNNYNDLDEPTAIRTLNLSKELGFSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGATNLESITLTGAQISSLPOTVCNQLPNLQVLDLS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGATNLESITLTGAQISSLPOTVCNQLPNLQVLDLS 360
QY 361 YNLLEDLPSFSVCKLOKIDLRHNEYIEIKVDTFOQLLSRLSLNLANWKIATHNAST 420
DB 361 YNLLEDLPSFSVCKLOKIDLRHNEYIEIKVDTFOQLLSRLSLNLANWKIATHNAST 420
QY 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIMPYAYOCC 480
DB 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIMPYAYOCC 480
QY 481 AFGVCENAYKISQNMKNDSMDLHKDAGMFAQDERDLEDFLLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISQNMKNDSMDLHKDAGMFAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPFKPCBHLDDGMLIRIGWTTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
DB 541 CSPSPGPFKPCBHLDDGMLIRIGWTTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600

QY 601 AVNMLTGVSASVLAGVDATFGSPARHGAWENGVGCHVIGFSLTFASSSVFLTLAAL 660
DB 601 AVNMLTGVSASVLAGVDATFGSPARHGAWENGVGCHVIGFSLTFASSSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLALTAAVPLLGSGSKYGASPLCLPLPFGEP 720
DB 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLALTAAVPLLGSGSKYGASPLCLPLPFGEP 720
QY 721 STMGYWVALILNLSLFLMNTIATKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYWVALILNLSLFLMNTIATKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFISFSSSLNLTISPVEVINKFILLVVVPLPACLNPLLYILFNHFHFKEDLVSLRKQTV 840
DB 781 PVAFISFSSSLNLTISPVEVINKFILLVVVPLPACLNPLLYILFNHFHFKEDLVSLRKQTV 840
QY 841 WTRSRKHPSLMSINSDDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
DB 841 WTRSRKHPSLMSINSDDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
DB 901 VAFVPECL 907
RESULT 10
ADN40013
ID ADN40013 standard; protein; 907 AA.
XX
XX AC ADN40013;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C383.
XX DE Human; differential expression; cancer; angiogenic disorder;
XX KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX KW inflammatory disease; autoimmune disease;
XX KW retinal neovascularisation syndrome; scarring; uterine fibroid;
XX KW detection; diagnosis; prognosis; drug screening; drug targeting;
XX KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-035250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-036809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-039775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39796.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
PS
PS Claim 12; SEQ ID NO C383; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
XX Sequence 907 AA;
SQ

Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRLRVDCSLGLSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRLRVDCSLGLSEL 60
Qy 61 PSNLSVFTSYLDLSMNNISQLPNPLPSRLFLEELRAGNALTYTPKGAFTGLSKVLK 120
Db 61 PSNLSVFTSYLDLSMNNISQLPNPLPSRLFLEELRAGNALTYTPKGAFTGLSKVLK 120
Qy 121 LQNNQLRHVPTEALQNLRSLSQLRLDANHISYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQLRLDANHISYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Qy 181 AFRSLALQAMTLALNKTHIPDYAFGNLSLVLLHNNRTHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLALQAMTLALNKTHIPDYAFGNLSLVLLHNNRTHSLGKKCFDGLHSLETLD 240
Qy 241 LNYNNLDEFPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Db 241 LNYNNLDEFPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Qy 301 FOHLPELRTLINGASQITEPPDLTGTANLSLTLTGAQISLSPQTVCNQLPNQVLDSL 360
Db 301 FOHLPELRTLINGASQITEPPDLTGTANLSLTLTGAQISLSPQTVCNQLPNQVLDSL 360
Qy 361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDFTFOQLLSLRLNLANWKIAIHPNAPST 420
Db 361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDFTFOQLLSLRLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSSNLLSFFITGLHGLTHLKTGNHALQSLISENPELKVIEPVAQCC 480
Db 421 LPSLIKLDLSSNLLSFFITGLHGLTHLKTGNHALQSLISENPELKVIEPVAQCC 480
Qy 481 AFGVCENAYKISQNWKGDNSMDLHKDKAGMFOQDERDLEDLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKGDNSMDLHKDKAGMFOQDERDLEDLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPFKCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600

Db 541 CSPSPGPFKCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLAVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Qy 661 ERGFSVKYSKAFETKAPSSLSKVIILLCALALTAALVPLLGSSKYGASPLCLPLPGCEP 720
Db 661 ERGFSVKYSKAFETKAPSSLSKVIILLCALALTAALVPLLGSSKYGASPLCLPLPGCEP 720
Qy 721 STMGYVALIILNLSLCFLMTIATYKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNLSLCFLMTIATYKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSLINLTFISPEVVKFILLVVVPLPACINPLLYLFLNPHPKEDLVSRKQTYV 840
Db 781 PVAFLSPSSLINLTFISPEVVKFILLVVVPLPACINPLLYLFLNPHPKEDLVSRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITYDLPPSPVPSPAYPVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITYDLPPSPVPSPAYPVTTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907
RESULT 11
ADN39531
ID ADN39531 standard; protein; 907 AA.
XX
AC ADN39531;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A131.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX

PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;	
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zielenk A;	
XX		
DR	WPI: 2003-468649/44.	
DR	N-PSDB; ADN39530.	
XX		
PT	Determining the presence or absence of a pathological cell in a patient,	
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting	
PT	a nucleic acid in a biological sample.	
XX		
PS	Claim 12; SEQ ID NO A131; 1385pp; English.	
XX		
CC	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)	
CC	whose expression is upregulated or downregulated in specific cancers or	
CC	other diseases such as angiogenic or fibrotic disorders, and to methods	
CC	of determining the presence or absence of a pathological cell in a	
CC	patient by detecting a nucleic acid at least 80% identical to those of	
CC	the invention or by detecting a polypeptide of the invention. The	
CC	invention also relates to expression vectors and host cells comprising a	
CC	nucleic acid of the invention; antibodies which specifically bind a	
CC	polypeptide of the invention; use of such antibodies for drug targeting;	
CC	and methods of screening for modulators of activity or expression of the	
CC	polypeptides and nucleic acids. The nucleic acids, polypeptides,	
CC	antibodies and methods are useful for diagnosing, prognosing and treating	
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,	
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal	
CC	neovascularisation syndromes, scarring and uterine fibroids. They may	
CC	also be useful in wound healing and in contraception. The present	
CC	sequence represents a polypeptide of the invention.	
XX		
SQ	Sequence 907 AA;	
	Query Match	
	Best Local Similarity 100.0%; Score 907; DB 7; Length 907;	
	Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60	
Db	1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60	
Qy	61 PSNLSVFTSYLDLSNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLSLKVL 120	
Db	61 PSNLSVFTSYLDLSNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLSLKVL 120	
Qy	121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVPPSCFSGLSLRLHLDNALTETIPVQ 180	
Db	121 LQNNQLRHVPTEALQNLRSQSLRDANHSIVPPSCFSGLSLRLHLDNALTETIPVQ 180	
Qy	181 AFRSLALQAMTLALNKIHHIPDYAFGNLSSLVVLHNNRIHSLGKKCFDGLHSLETLD 240	
Db	181 AFRSLALQAMTLALNKIHHIPDYAFGNLSSLVVLHNNRIHSLGKKCFDGLHSLETLD 240	
Qy	241 LNNNLDEFPFAIRTLNKLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300	
Db	241 LNNNLDEFPFAIRTLNKLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300	
Qy	301 FOHLPELRTLNGASQITTEPDLTGTLNLSLTGTGAQISLPTVCNQLPNLQVLDLS 360	
Db	301 FOHLPELRTLNGASQITTEPDLTGTLNLSLTGTGAQISLPTVCNQLPNLQVLDLS 360	
Qy	361 YNLLEDLPSFVSCQKQLKIDLRHNEIYKVDVTFQQLSLRLSLNLANWKIATIHNAFST 420	
Db	361 YNLLEDLPSFVSCQKQLKIDLRHNEIYKVDVTFQQLSLRLSLNLANWKIATIHNAFST 420	
Qy	421 LPSLIKLDLSNLSLSPFITGLHGLTHLKLGNHALQSLISSENPELKIEMPVAYQCC 480	
Db	421 LPSLIKLDLSNLSLSPFITGLHGLTHLKLGNHALQSLISSENPELKIEMPVAYQCC 480	
Qy	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOQDERDLELDFEEDLKALHSVQ 540	
Db	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOQDERDLELDFEEDLKALHSVQ 540	

Qy	541 CSPSPGPPKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600	
Db	541 CSPSPGPPKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600	
Qy	601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFLSTFASSESVFLTLTAA 660	
Db	601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWENGVCVIGFLSTFASSESVFLTLTAA 660	
Qy	661 ERGSVKYSKAFETKAPFSSSLKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGEP 720	
Db	661 ERGSVKYSKAFETKAPFSSSLKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGEP 720	
Qy	721 STMGPMVALIILNSLCFLMMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780	
Db	721 STMGPMVALIILNSLCFLMMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780	
Qy	781 PVAFSLFSSLLNLFISPEVKIFLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTV 840	
Db	781 PVAFSLFSSLLNLFISPEVKIFLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTV 840	
Qy	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPSPSSVPSPAYPVTESCHLSS 900	
Db	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPSPSSVPSPAYPVTESCHLSS 900	
Qy	901 VAFVPC 907	
Db	901 VAFVPC 907	

RESULT 12

ADN39628

ID ADN39628 standard; protein; 907 AA.

XX

AC ADN39628;

XX

DT 17-JUN-2004 (first entry)

XX

Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A228.

DE

Human; differential expression; cancer; angiogenic disorder;

KW

fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;

KW

inflammatory disease; autoimmune disease;

KW

retinal neovascularisation syndrome; scarring; uterine fibroid;

KW

detection; diagnosis; prognosis; drug screening; drug targeting;

KW

wound healing; contraception; cytostatic; cardiant; immunomodulatory;

KW

vulnary; gene therapy; vaccine.

XX

Homo sapiens.

XX

WO2003042661-A2.

PN

22-MAY-2003.

PD

13-NOV-2002; 2002WO-US036810.

XX

13-NOV-2001; 2001US-0350666P.

PR

21-NOV-2001; 2001US-0332464P.

PR

29-NOV-2001; 2001US-0334393P.

PR

03-DEC-2001; 2001US-0335394P.

PR

14-DEC-2001; 2001US-0340376P.

PR

08-JAN-2002; 2002US-0347211P.

PR

10-JAN-2002; 2002US-0347349P.

PR

08-FEB-2002; 2002US-035250P.

PR

13-FEB-2002; 2002US-0356714P.

PR

20-FEB-2002; 2002US-0359077P.

PR

29-MAR-2002; 2002US-0368809P.

PR

04-APR-2002; 2002US-0370110P.

PR

12-APR-2002; 2002US-0372246P.

PR

05-JUN-2002; 2002US-0386614P.

PR

16-JUL-2002; 2002US-0396839P.

PR

22-JUL-2002; 2002US-0397775P.

PR

22-JUL-2002; 2002US-0397845P.

```
PR 09-SEP-2002; 2002US-0409450P.
XX (ROSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39627.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO A228; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGGRMLLRVDCSDLGLSEL 60
DB 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGGRMLLRVDCSDLGLSEL 60
QY 61 PSNLVFTSYLDLSNNISQLLPNLPFLRFLRFLRAGNALTYPKGAFTGLYSLKVLV 120
DB 61 PSNLVFTSYLDLSNNISQLLPNLPFLRFLRFLRAGNALTYPKGAFTGLYSLKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
QY 181 AFRSLALQAMTALANKIHHPDYAFGNLSLVVLHNNRHISLGGKCFDGLHSLTLD 240
DB 181 AFRSLALQAMTALANKIHHPDYAFGNLSLVVLHNNRHISLGGKCFDGLHSLTLD 240
QY 241 LMYNNLDEFPPTAIRLSNLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
DB 241 LMYNNLDEFPPTAIRLSNLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRLTLUNGASQITTEPDLTGTLNLESLSLTGAQISSLPQVCNQLPNLVLDLS 360
DB 301 FOHLPELRLTLUNGASQITTEPDLTGTLNLESLSLTGAQISSLPQVCNQLPNLVLDLS 360
QY 361 YNLLDLPSFSVCQKLOKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAST 420
DB 361 YNLLDLPSFSVCQKLOKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAST 420
QY 421 LPSLKLDSLNNLSLSPFITGLHGLTHLKTGNHALQSLISSENPELKIEMPIAYOCC 480
DB 421 LPSLKLDSLNNLSLSPFITGLHGLTHLKTGNHALQSLISSENPELKIEMPIAYOCC 480
QY 481 AFGVCENAYKISQNNKGNSSMDLHKDAGMFOAQDERDLEDLFDLDFEEDLKALHSVQ 540
```

```
DB 481 AFGVCENAYKISQNNKGNSSMDLHKDAGMFOAQDERDLEDLFDLDFEEDLKALHSVQ 540
QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSAAVLGADAFPGSFARHCAMWENGVCCHVIGFLSTIFASESSVELLTLAAL 660
DB 601 AVNMLTGVSAAVLGADAFPGSFARHCAMWENGVCCHVIGFLSTIFASESSVELLTLAAL 660
QY 661 ERGSFVKYSKAFETKAPFSSLKVIILLCALIALTMAAVPLLGSKYKASPLCLPFGCEP 720
DB 661 ERGSFVKYSKAFETKAPFSSLKVIILLCALIALTMAAVPLLGSKYKASPLCLPFGCEP 720
QY 721 STMGYVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENINDCSMVKHIALLLFNLCILNC 780
DB 721 STMGYVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENINDCSMVKHIALLLFNLCILNC 780
QY 781 PVAFLSFSSILNLTPIISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQIV 840
DB 781 PVAFLSFSSILNLTPIISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQIV 840
QY 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSSITVDLPSPSPVPAYPVFTESCHLSS 900
DB 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSSITVDLPSPSPVPAYPVFTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

RESULT 13
ADO29408
ID ADO29408 standard; protein; 907 AA.
XX AC ADO29408;
XX DT 29-JUL-2004 (first entry)
XX DE Human GPCR GPR49, SEQ ID NO:510.
XX KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
KW cytostatic; antinflammatory; vasotropic; antiangiinal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW viricide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW receptor.
XX OS Homo sapiens.
XX PN WO2004040000-A2.
XX PD 13-MAY-2004.
XX PF 09-SEP-2003; 2003WO-US028226.
XX PR 09-SEP-2002; 2002US-0409303P.
XX PR 09-APR-2003; 2003US-0461329P.
XX PA (PRIM-) PRIMAL INC.
XX PI Gaitanaris GA, Bergmann JF, Gragerov A, Hohmann J, Li P;
PI Madisen L, Mcilwain KL, Pavlova MN, Vassilatis D, Zeng H;
```

XX WPI; 2004-390329/36.
DR N-PSDB; ADO29923.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 510; 542pp; English.
XX
XX The invention relates to human and mouse G protein-coupled receptors
XX (GPCRs) and nucleic acids encoding them. The invention also relates to
XX sequences at least 90% identical to the GPCR proteins and nucleic acids
XX of the invention; methods of treating, preventing or diagnosing diseases
XX associated with GPCRs of the invention; methods of screening for
XX compounds useful in the treatment of GPCR-related diseases; a transgenic
XX mouse comprising a GPCR gene of the invention; a mouse comprising a
XX mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
XX from the transgenic mice; kits comprising several mice, each of which has
XX a mutation in a different GPCR gene of the invention; and kits comprising
XX probes which hybridize to GPCR polynucleotides of the invention. The
XX invention further discloses variants of the GPCR polypeptides and vectors
XX comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
XX be used in the diagnosis, treatment or prevention of a wide variety of
XX diseases including neurological disorders (e.g., Alzheimer's disease,
XX depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
XX disorders of the adrenal gland; disorders of the colon or intestine
XX (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
XX syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
XX myocardial infarction); immune disorders (e.g., autoimmune disorders or
XX anaemia or leukaemia); and disorders (e.g., osteoarthritis, rheumatoid
XX arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
XX obesity, enzyme deficiency-related diseases or vitamin deficiency-related
XX diseases); and disorders of the kidney, liver, lung, breast, ovary,
XX uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
XX thyroid (e.g., cancers). The present sequence represents a GPCR of the
XX invention. Note: The full sequence data for this patent did not form part
XX of the printed specification; those sequences not shown were obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 907; DB 8; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PNLVSFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYTPKGAFTGLYSKVLUM 120
DB 61 PNLVSFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYTPKGAFTGLYSKVLUM 120
QY 121 LQNNQLRHVPTALQNLRLQSLRLDANHSVVPSCFSLHSLRHLMDNLTALTEIPVQ 180
DB 121 LQNNQLRHVPTALQNLRLQSLRLDANHSVVPSCFSLHSLRHLMDNLTALTEIPVQ 180
QY 181 AFRSLAQMTALNKLHIIHIDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLTLD 240
DB 181 AFRSLAQMTALNKLHIIHIDYAFGNLSLVLLHNNRIHSLGKKCFDGLHSLTLD 240
QY 241 LNNYNDLDEPTAIRLTSNKLGFHNSNIRTSIPKAFVGNPSLTIHFYDNPQVGRSA 300
DB 241 LNNYNDLDEPTAIRLTSNKLGFHNSNIRTSIPKAFVGNPSLTIHFYDNPQVGRSA 300
QY 301 FOHLPELRLTLNGASQITFEFPLDTGTANLESILTGAQISLPQVCNQLPNQLVDLS 360
DB 301 FOHLPELRLTLNGASQITFEFPLDTGTANLESILTGAQISLPQVCNQLPNQLVDLS 360

QY 361 YNLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHPNAST 420
DB 361 YNLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHPNAST 420
QY 421 LPSLTKLDLSSNLSFPFITGLHGLTHLKTGNHALQSLISSENPELKVIEPIYAYQCC 480
DB 421 LPSLTKLDLSSNLSFPFITGLHGLTHLKTGNHALQSLISSENPELKVIEPIYAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGDNSMDDLHKKQAGMFOAQDERDLEDFLDDFEEDKALHVSQ 540
DB 481 AFGVCENAYKISQNWKNKGDNSMDDLHKKQAGMFOAQDERDLEDFLDDFEEDKALHVSQ 540
QY 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLADFTFGSFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVAVLADFTFGSFARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLALTMAAVPLLGSKYKYGASPLCLPLPFGP 720
DB 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLALTMAAVPLLGSKYKYGASPLCLPLPFGP 720
QY 721 STMGYVVALIILNSLCFLMMTIAVTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYVVALIILNSLCFLMMTIAVTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFSLSPSLNLTFTISPEVIKFIILVVVPLPACLNPLLYILFNPHEKEDLSLRKQTVV 840
DB 781 PVAFSLSPSLNLTFTISPEVIKFIILVVVPLPACLNPLLYILFNPHEKEDLSLRKQTVV 840
QY 841 WTRSKHPSLMSINSDVVKQSCDSTQALVFTSSITVDLPSSVSPSPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDVVKQSCDSTQALVFTSSITVDLPSSVSPSPAYPVTESCHLSS 900
QY 901 VAFVPCCL 907
DB 901 VAFVPCCL 907
RESULT 14
ADQ80369
ID ADQ80369 standard; protein; 907 AA.
XX
XX AC ADQ80369;
XX
XX DT 21-OCT-2004 (first entry)
XX
XX DE G protein-coupled receptor 49 protein.
XX
XX KW cytostatic; epidermal growth factor receptor modulator; identification;
XX therapeutic response; cancer; EGFR; biomarker.
XX
XX OS Homo sapiens.
XX
XX PN WO2004063709-A2.
XX
XX PD 29-JUL-2004.
XX
XX PF 08-JAN-2004; 2004WO-US000368.
XX
XX PR 08-JAN-2003; 2003US-0438735P.
XX
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX PI Amler LC, Januario T;
XX
XX DR WPI; 2004-544114/52.
XX
XX DR N-PSDB; ADQ80249.
XX
XX PT Identifying a mammal that will respond therapeutically to a method of
XX treating cancer comprises comparing the level of a biomarker in a mammal

PT before and after exposure to an epidermal growth factor receptor (EGFR)
PT modulator.
PS Disclosure; SEQ ID NO 141; 520pp; English.
XX The invention relates to a method of identifying a mammal that will
CC respond therapeutically to a method of treating cancer by administering
CC an epidermal growth factor receptor (EGFR) modulator by comparing the
CC level of a biomarker in a mammal before and after exposure to an EGFR
CC modulator. The method comprises: (a) measuring, in the mammal, the level
CC of at least one biomarker identified in the specification; (b) exposing
CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
CC level of the biomarker, where a difference in the level in step (c)
CC compared to step (a) indicates that the mammal will respond
CC therapeutically to the method of treating cancer. The method and
CC biomarkers are useful for identifying a mammal that will respond
CC therapeutically to a method of treating cancer by administering an
CC epidermal growth factor receptor (EGFR) modulator. This sequence
CC corresponds to one of the biomarkers whose levels of expression is
CC measured in the method of the invention.
XX
XX Sequence 907 AA;
Query Match 100.0%; Score 907; DB 8; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSDIGLSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSDIGLSEL 60
QY 61 PSLNSVFTSYDLSMNNISQLLPNLSIRFLFEEURLAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSLNSVFTSYDLSMNNISQLLPNLSIRFLFEEURLAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVPPSCFSGLSLHRLMLDNLALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLSLQSLRLDANHSIVPPSCFSGLSLHRLMLDNLALTEIPVQ 180
QY 181 AFRSLALQAMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKPGLHSLETLTD 240
Db 181 AFRSLALQAMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKPGLHSLETLTD 240
QY 241 LNNYNNLDEFFPAIRTLNLSKELGFHNNIRSIPEKAFVGNPNSLTIHFYDNPFIQVGRSA 300
Db 241 LNNYNNLDEFFPAIRTLNLSKELGFHNNIRSIPEKAFVGNPNSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGATNLESLLTGAQISLSPQTVCNQPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITEFPDLTGATNLESLLTGAQISLSPQTVCNQPNLQVLDLS 360
QY 361 YNLLLEDPSFVSCQKQKIDLRHNEIYEIKVDTFQQLLSLRSLNLANWKIAIHHNAPST 420
Db 361 YNLLLEDPSFVSCQKQKIDLRHNEIYEIKVDTFQQLLSLRSLNLANWKIAIHHNAPST 420
QY 421 LPSLTKLDLSSNLSSEPTGLGHITLKTGNHALQSLISENPELKIEMPVAYOCC 480
Db 421 LPSLTKLDLSSNLSSEPTGLGHITLKTGNHALQSLISENPELKIEMPVAYOCC 480
QY 481 AFGVCENAYKISQNNKGNDSMDLHKDDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNNKGNDSMDLHKDDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYSIPKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYSIPKLLIGVIA 600
QY 601 AVNMLTGVSASVAGVDAFTGSGFARHGAWENGVCVIGPLSIPASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVAGVDAFTGSGFARHGAWENGVCVIGPLSIPASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPFSSSLKVIILLCALLALTWAAVPLLGSSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSSLKVIILLCALLALTWAAVPLLGSSKYGASPLCLPLPFGEP 720

Db 661 ERGFSVKYSAKFETKAPFSSSLKVIILLCALLALTWAAVPLLGSSKYGASPLCLPLPFGEP 720
QY 721 STMGYVALIILNLSLCFLMMTIATKLYCNLDKGDLENINWDCSMVKHTALLFTNCILNC 780
Db 721 STMGYVALIILNLSLCFLMMTIATKLYCNLDKGDLENINWDCSMVKHTALLFTNCILNC 780
QY 781 PVAFLPSSLLNLFIISPEVIFKILLVVVPLPACINPLLYILFNPFKEDLVSLRKQTYV 840
Db 781 PVAFLPSSLLNLFIISPEVIFKILLVVVPLPACINPLLYILFNPFKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSPSPVPAYPVTSSCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSPSPVPAYPVTSSCHLSS 900
QY 901 VAFVPCPL 907
Db 901 VAFVPCPL 907
RESULT 15
ADR67868
ID ADR67868 standard; protein; 907 AA.
XX
AC ADR67868;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human HG38 protein.
XX
KW human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
XX
OS Homo sapiens.
XX
PN WO2004074436-A2.
XX
PD 02-SEP-2004.
XX
PF 11-FEB-2004; 2004WO-US004060.
XX
PR 19-FEB-2003; 2003US-0448959P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Lasek AW;
XX
PI WPI; 2004-652946/63.
DR N-PSDB; ADR67869.
DR
PT Detecting colon or lung cancer, by determining amount of protein in
PT sample, comparing amount of protein to standard, and differential
PT expression of protein in sample indicates colon or lung cancer.
XX
XX Claim 4; SEQ ID NO 1; 79pp; English.
XX This sequence represents the human G protein-coupled receptor (GPCR)
XX known as HG38. The HG38 protein and corresponding nucleic acid, may be
XX used in the method of the invention for detecting colon or lung cancer.
XX The first method involves performing an assay to determine the amount of
XX HG38 in a sample of colon or lung tissue, and comparing the amount of
XX protein to standard, thus detecting expression of protein in sample,
XX where differential expression of protein in sample when compared with the
XX standard is diagnostic of colon or lung cancer. The second method
XX involves hybridizing a composition comprising the HG38 coding sequence,
XX or its complement, and a labelling moiety, to nucleic acids of a sample
XX of colon or lung tissue under conditions to form at least one
XX hybridization complex, detecting hybridization complex formation, and
XX comparing complex formation to a standard, where the comparison reflects
XX differential expression of the polynucleotide in the sample relative to
XX the standard and is diagnostic of a colon or lung cancer. This first
XX method enables earlier diagnosis before the patient is symptomatic. Anti-
XX HG38 antibodies are useful for treating colon or lung cancer.
XX Sequence 907 AA;

Query Match	100.0%;	Score 907;	DB 8;	Length 907;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 907;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLGSEL	60	
DB	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLGSEL	60	
QY	61	PSNLSVFTSYLDLSMNNISQLPNLPSLRPLEELRLAGNALTYIPKGAFTGLYSKVLV	120	
DB	61	PSNLSVFTSYLDLSMNNISQLPNLPSLRPLEELRLAGNALTYIPKGAFTGLYSKVLV	120	
QY	121	LQNNQLRHVPTALQNLRLSLSLQSLRLDANHSYVPPSCFSGLSRLHMLDDNALTEIPVQ	180	
DB	121	LQNNQLRHVPTALQNLRLSLSLQSLRLDANHSYVPPSCFSGLSRLHMLDDNALTEIPVQ	180	
QY	181	AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240	
DB	181	AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240	
QY	241	LNYYNLDREFFPAIRTLNKLGFHNSNIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA	300	
DB	241	LNYYNLDREFFPAIRTLNKLGFHNSNIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA	300	
QY	301	FQHLPELRTLTLNGASQITEPPDLTG TANLES LITGAQISSLPQTCVNCNLPNLQVLDLS	360	
DB	301	FQHLPELRTLTLNGASQITEPPDLTG TANLES LITGAQISSLPQTCVNCNLPNLQVLDLS	360	
QY	361	YNLEEDLPSFVCQKQKIDLRHNEIYEIKVDTFQOQLSLRSLNLANWKIAIITHFNAPST	420	
DB	361	YNLEEDLPSFVCQKQKIDLRHNEIYEIKVDTFQOQLSLRSLNLANWKIAIITHFNAPST	420	
QY	421	LPSLILKDLSSNLLSFPITGLHGTTHLKTGNHALQSLISSENPELKVTEMPYAYOCC	480	
DB	421	LPSLILKDLSSNLLSFPITGLHGTTHLKTGNHALQSLISSENPELKVTEMPYAYOCC	480	
QY	481	AFGCENAYKISNOWNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540	
DB	481	AFGCENAYKISNOWNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540	
QY	541	CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600	
DB	541	CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600	
QY	601	AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASSESVFLLTLAAL	660	
DB	601	AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASSESVFLLTLAAL	660	
QY	661	ERGSVKYSAKPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP	720	
DB	661	ERGSVKYSAKPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP	720	
QY	721	STMGYWVALIILNLSICFLMWTIATYKLYCNLDKGDLENIDCSMVKHIALLFTNCILNC	780	
DB	721	STMGYWVALIILNLSICFLMWTIATYKLYCNLDKGDLENIDCSMVKHIALLFTNCILNC	780	
QY	781	PVAFLSFSSLNLTFTISPEVIKFI LLVVVPLPACLNPLLYILFNPHKEDLVSLRKQTYV	840	
DB	781	PVAFLSFSSLNLTFTISPEVIKFI LLVVVPLPACLNPLLYILFNPHKEDLVSLRKQTYV	840	
QY	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSSITYDLPSSVPSPAYPVTESCHLSS	900	
DB	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSSITYDLPSSVPSPAYPVTESCHLSS	900	
QY	901	VAFVPC L 907		
DB	901	VAFVPC L 907		

Search completed: July 12, 2005, 07:51:17
Job time : 112 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: July 12, 2005, 07:38:42 ; Search time 29 Seconds
(without alignments)
3009.262 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTESCHLSSVAFVPCFL 907

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	2 JG0176	orphan G protein-c
2	51	5.6	907	2 JG0193	G protein-coupled
3	9	1.0	195	2 AF2638	conserved hypothet
4	9	1.0	209	2 G97420	hypothetical prote
5	9	1.0	580	2 S06057	gene NDI intron 3
6	9	1.0	869	2 A71400	probable disease r
7	9	1.0	1978	2 S77257	hypothetical prote
8	8	0.9	162	2 F97440	hypothetical prote
9	8	0.9	178	2 A87605	Dps family protein
10	8	0.9	179	2 S75434	hypothetical prote
11	8	0.9	238	2 AD1368	amino acid ABC-tra
12	8	0.9	267	2 AF3511	homospermidine syn
13	8	0.9	293	2 B71686	RNA polymerase sig
14	8	0.9	299	2 S61248	hypothetical prote
15	8	0.9	299	2 A75591	hypothetical prote
16	8	0.9	324	2 T20679	hypothetical prote
17	8	0.9	329	2 H64143	hypothetical prote
18	8	0.9	402	2 T15490	hypothetical prote
19	8	0.9	431	2 H84392	O-acetyl homoserin
20	8	0.9	467	2 D86583	ubiquinone oxidore
21	8	0.9	467	2 E72040	probable sodium-tr
22	8	0.9	478	2 S75572	glucosyltransferas
23	8	0.9	481	2 AD3020	homospermidine syn
24	8	0.9	481	2 E98264	homospermidine syn
25	8	0.9	538	1 D64164	hypothetical prote
26	8	0.9	671	2 D84648	probable disease r
27	8	0.9	674	2 JG5104	transcription init
28	8	0.9	678	2 C97114	recG helicase [imp
29	8	0.9	729	2 F86308	Similar to disease

30	8	0.9	732	2 S25995	hypothetical prote
31	8	0.9	851	2 S67285	NUDI protein - yea
32	8	0.9	967	2 T48210	hypothetical prote
33	8	0.9	976	2 B84659	probable receptor-
34	8	0.9	980	2 T05414	protein kinase hom
35	8	0.9	1159	1 A44280	inner layer protei
36	8	0.9	1286	2 T33476	hypothetical prote
37	8	0.9	1614	2 T29861	hypothetical prote
38	8	0.9	2025	2 T03884	hypothetical prote
39	7	0.8	94	2 E86889	hypothetical prote
40	7	0.8	98	2 A58889	NADH2 dehydrogenas
41	7	0.8	103	2 AE3419	hypothetical prote
42	7	0.8	105	4 S59325	hypothetical prote
43	7	0.8	106	2 T04804	hypothetical prote
44	7	0.8	106	2 E83435	hypothetical prote
45	7	0.8	108	2 S75540	hypothetical prote

ALIGNMENTS

RESULT 1

JB0176

orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man)

C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004

C:Accession: JB0176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A>Title: Identification and cloning of an orphan G protein-coupled receptor of the glyc

A:Reference number: JB0176; MUID:98308104; PMID:9642114

A:Accession: JB0176

A:Molecule type: mRNA

A:Residues: 1-907 <MCD>

A:Cross-references: UNIPROT:O75473; GB:AF062006; NID:G3366801; PIDN:AAC28019,1; PID:G336

C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C:Genetics:

A:Gene: HG38

A:Map position: 12q22-23

F:1-21/Domain: signal sequence #status predicted <SIG>

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:633-660/Domain: transmembrane #status predicted <TM3>

F:681-701/Domain: transmembrane #status predicted <TM4>

F:725-744/Domain: transmembrane #status predicted <TM5>

F:768-791/Domain: transmembrane #status predicted <TM6>

F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 100.0%; Score 907; DB 2; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGSSPRSGVTLRGCPETHCHCPDGRMLRVDCSDGLSEL 60

Db 1 MDTSLRGVLLSLPVLQLATGSSPRSGVTLRGCPETHCHCPDGRMLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYLDLNNNISQLLPNPLPSLRFLRELRLAGNALTYIPKGAFTGLYSKVLV 120

Db 61 PSNLSVFTSYLDLNNNISQLLPNPLPSLRFLRELRLAGNALTYIPKGAFTGLYSKVLV 120

Qy 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSPCSGLHSRLHLDLNNALTEIPVQ 180

Db 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHI SYVPPSPCSGLHSRLHLDLNNALTEIPVQ 180

Qy 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFGLHSLFTLD 240

Db 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFGLHSLFTLD 240

Qy 241 LNNYNLDEFFTAIRTLNKLGLFHSNNIRIPEKAFVGNPSLTIHFYDNP IQVGRSA 300

Db 241 LNNYNLDEFFTAIRTLNKLGLFHSNNIRIPEKAFVGNPSLTIHFYDNP IQVGRSA 300

Qy 301 FOHLPELRLTLNGASQITTEPFDLTGTANLES LTLTGAQISLSPQTVCNQLPNLQVLDLS 360

Db	301	FOHLEPELTLTLNGASQITEPDLTGTANLESLLTITGAIQISLPTVCNQPLNQLVLDLS	360
Qy	361	YNLEBDFSFVSCQKLOKIDLRHNEIYBIKVDTPQOLLSLRSLNLAWNKIAIHHNAPST	420
Db	361	YNLEBDFSFVSCQKLOKIDLRHNEIYBIKVDTPQOLLSLRSLNLAWNKIAIHHNAPST	420
Qy	421	LPSLIKIDLSNLLSPFITGLHGLTHLKTGNHALQSLISENPPELKVIEMPHYAYOCC	480
Db	421	LPSLIKIDLSNLLSPFITGLHGLTHLKTGNHALQSLISENPPELKVIEMPHYAYOCC	480
Qy	481	AFVCENAYKISQNWKNKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ	540
Db	481	AFVCENAYKISQNWKNKDNSSMDLHKDAGMFOQDERDLEDFLLDFEEDLKALHSVQ	540
Qy	541	CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
Db	541	CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
Qy	601	AVNMLTGVSSAVLAGVDATFGSPARHGAWENGVCVIGFLSIFASESSVFLITLAAL	660
Db	601	AVNMLTGVSSAVLAGVDATFGSPARHGAWENGVCVIGFLSIFASESSVFLITLAAL	660
Qy	661	ERGFVKYSAPETKAPFSSLKVIIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEF	720
Db	661	ERGFVKYSAPETKAPFSSLKVIIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEF	720
Qy	721	STMGYMVALIILNSLCFLMWTIATKLYCNLDKGDLENIDCSMVVKHIALLLFTNCILNC	780
Db	721	STMGYMVALIILNSLCFLMWTIATKLYCNLDKGDLENIDCSMVVKHIALLLFTNCILNC	780
Qy	781	PVAFLSFSLNLTIFISBEVVKFILLVVVPLPACLNPLLYLILFNHPKEDLVSLRKQTYV	840
Db	781	PVAFLSFSLNLTIFISBEVVKFILLVVVPLPACLNPLLYLILFNHPKEDLVSLRKQTYV	840
Qy	841	WTRSKHPLMSINDSDVKQSCDSTQALVFTFTSSITVDLPSSVPSPAYPVPTESCHLSS	900
Db	841	WTRSKHPLMSINDSDVKQSCDSTQALVFTFTSSITVDLPSSVPSPAYPVPTESCHLSS	900
Qy	901	VAFVPCCL 907	
Db	901	VAFVPCCL 907	
RESULT 2			
JG0193			
G protein-coupled receptor FXR - mouse			
C/Species: Mus musculus (house mouse)			
C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000			
C/Accession: JG0193			
R/Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.			
Biochem. Biophys. Res. Commun. 254, 273-279, 1999			
A/Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin A			
A/Reference number: JG0193; MUID:99121227; PMID:9920770			
A/Accession: JG0193			
A/Status: preliminary			
A/Molecule type: mRNA			
A/Residues: 1-907 <HER>			
Query Match			
Best Local Similarity 5.6%; Score 51; DB 2; Length 907;			
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	203	DYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLETLDLNNYNNLDEFFPTAI	253
Db	203	DYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLETLDLNNYNNLDEFFPTAI	253
RESULT 3			
AF2638			
conserved hypothetical protein Atu0506 [imported] - Agrobacterium tumefaciens (strain C58)			
C/Species: Agrobacterium tumefaciens			
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004			

C/Accession: AF2638			
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.			
; Karp, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.			
Science 294, 2317-2323, 2001			
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.E.W.			
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.			
A/Reference number: AB2577; MUID:21608550; PMID:11743193			
A/Accession: AF2638			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-195 <KUR>			
A/Cross-references: UNIPROT:Q8UHZ5; GB:AB008688; PIDN:AAL41524.1; PID:gi7738854; GSPDB:G			
A/Experimental source: strain C58 (Dupont)			
C/Genetics:			
A/Gene: Atu0506			
A/Map position: circular chromosome			
Query Match			
Best Local Similarity 1.0%; Score 9; DB 2; Length 195;			
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	25	PRSGVLLRG 33	
Db	145	PRSGVLLRG 153	
RESULT 4			
G97420			
hypothetical protein AGR_C_894 [imported] - Agrobacterium tumefaciens (strain C58, Cerezo)			
C/Species: Agrobacterium tumefaciens			
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004			
C/Accession: G97420			
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;			
Science 294, 2323-2328, 2001			
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens			
A/Reference number: A97359; MUID:21608551; PMID:11743194			
A/Accession: G97420			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-209 <KUR>			
A/Cross-references: UNIPROT:Q8UHZ5; GB:AE007869; PIDN:AAK86320.1; PID:gi15155439; GSPDB:G			
C/Genetics:			
A/Gene: AGR_C_894			
A/Map position: circular chromosome			
Query Match			
Best Local Similarity 1.0%; Score 9; DB 2; Length 209;			
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	25	PRSGVLLRG 33	
Db	159	PRSGVLLRG 167	
RESULT 5			
S06057			
gene ND1 intron 3 protein 2 - Podospora anserina mitochondrion			
C/Species: mitochondrion Podospora anserina			
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004			
C/Accession: S06057			
R/Cummings, D.J.; Domenico, J.M.; Michel, F.			
Curr. Genet. 14, 253-264, 1988			
A/Title: DNA sequence and organization of the mitochondrial ND1 gene from Podospora anserina			
A/Reference number: S06056; MUID:89063443; PMID:3197134			
A/Accession: S06057			
A/Molecule type: DNA			
A/Residues: 1-580 <CUM>			
A/Cross-references: UNIPROT:Q02714; EMBL:X13164			
C/Genetics:			
A/Genome: mitochondrion			

A;Genetic code: SGC3
C;Keywords: mitochondrion

Query Match 1.0%; Score 9; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 530 BEDLKALHS 538
|||||
DB 108 BEDLKALHS 116

RESULT 6

A71400
probable disease resistance protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: A71400

R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ana
C.; Chalwatzis, N.

A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215

A;Accession: A71400

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-869 <BEV>

A;Cross-references: UNIPROT:Q23253; GB:Z97335; NID:G2244747; PID:E326857; PID:G2244748
C;Genetics:

A;Map position: 4COP9-4G3845

Query Match 1.0%; Score 9; DB 2; Length 869;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 LPSLIKLDL 429
|||||
DB 334 LPSLIKLDL 342

RESULT 7

S77257
hypothetical protein sll1265 - Synechocystis sp. (strain PCC 6803)

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: S77257

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S77257

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1978 <KAN>

A;Cross-references: UNIPROT:P73551; EMBL:D90907; GB:AB001339; NID:G1652618; PIDN:BAAL759
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 1.0%; Score 9; DB 2; Length 1978;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 LTLTGAQIS 341
|||||
DB 526 LTLTGAQIS 534

RESULT 8

F97440

hypothetical protein AGR_C_1201 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: F97440

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman, B.;
A.; Liu, F.; Wollam, C.; Allinger, M.; Boughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: F97440

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-162 <KUR>

A;Cross-references: UNIPROT:Q8U5I8; GB:AE007869; PIDN:AAK86479.1; PID:G15155627; GSPDB:G

C;Genetics:

A;Gene: AGR_C_1201

A;Map position: circular chromosome

Query Match 0.9%; Score 8; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 LLALTMAA 697
|||||

DB 22 LLALTMAA 29

RESULT 9

A87605

Dps family protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: A87605

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87605

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-178 <STO>

A;Cross-references: UNIPROT:Q9A4G1; GB:AE005673; NID:G13424489; PIDN:AAK24837.1; GSPDB:G

C;Genetics:

A;Gene: CC2873

Query Match 0.9%; Score 8; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 AFGMLSSL 212
|||||

DB 104 AFGMLSSL 111

RESULT 10

S75434

hypothetical protein c05017 - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004

C;Accession: S75434

R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996

A;Title: Organizational characteristics and information content of an archaeal genome: I
A;Reference number: S73076; MUID:97055432; PMID:8899719

A;Accession: S75434

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-179 <SEN>

A;Cross-references: UNIPROT:P96006; EMBL:Y08257; NID:gl707772; PID:c283869; PID:gl707840
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Genetics:
A;Start codon: GTG

Query Match 0.9%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 679 SSLKVIIL 686
|||
Db 14 SSLKVIIL 21

RESULT 11
AD1368
A;Title: amino acid ABC-transporter (permease) homolog lmo2348 [imported] - *Listeria monocytogenes*
C;Species: *Listeria monocytogenes*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD1368
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluecker, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1368
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <GLA>
A;Cross-references: UNIPROT:Q8Y4T4; GB:NC_003210; PIDN:CAD00426.1; PID:gl6411836; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2348
C;Superfamily: histidine permease protein M

Query Match 0.9%; Score 8; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGVLLSLP 13
|||
Db 38 LGVLLSLP 45

RESULT 12
AF3511
A;Title: homoserimidine synthase [imported] - *Brucella melitensis* (strain 16M)
C;Species: *Brucella melitensis*
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AF3511
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesha, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A;Reference number: AD3252; PMID:11756688
A;Accession: AF3511
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <KUR>
A;Cross-references: UNIPROT:Q8Y502; GB:AE008918; PIDN:AAL53257.1; PID:gl7984137; GSPDB:C
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10016
A;Map position: II

Query Match 0.9%; Score 8; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 608 VSSAVLAG 615
|||||
Db 191 VSSAVLAG 198

RESULT 13

B71686
RNA polymerase sigma-32 factor (rpoH) RP303 - *Rickettsia prowazekii*

C;Species: *Rickettsia prowazekii*

C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: B71686

R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U.

Nature 396, 133-140, 1998

A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.

A;Reference number: A71630; MUID:99039499; PMID:9823893

A;Accession: B71686

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-293 <AND>

A;Cross-references: UNIPROT:Q9ZDM4; GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAAL476;

A;Experimental source: strain Madrid E

C;Genetics:

A;Gene: rpoH; RP303

C;Superfamily: transcription initiation factor sigma katF; transcription initiation factor

F;55-287/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match 0.9%; Score 8; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 851 SINSDDVE 858
|||||
Db 190 SINSDDVE 197

RESULT 14

S61248
hypothetical protein - bovine herpesvirus 1

C;Species: bovine herpesvirus 1

C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004

C;Accession: S61248

R;Vicek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Sch

submitted to the EMBL Data Library, January 1995

A;Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus

A;Reference number: S61233

A;Accession: S61248

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-299 <VLC>

A;Cross-references: UNIPROT:Q65574; EMBL:Z48053; NID:g971311; PIDN:CAA88126.1; PID:g9713

C;Superfamily: varicella-zoster virus gene 53 protein

Query Match 0.9%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 TPTSSSIT 877
|||||
Db 60 TPTSSSIT 67

RESULT 15

A75591
hypothetical protein - *Deinococcus radiodurans* (strain R1)

C;Species: *Deinococcus radiodurans*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: A75591

R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I

S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mai

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: A75591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <WHI>
A;Cross-references: UNIPROT:O9RYG0; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF1250
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0358
A;Map position: 2

Query Match 0.9%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 SQQLEPNPL 86
|||
Db 66 SQQLEPNPL 73

Search completed: July 12, 2005, 07:53:43
Job time : 35 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:37:57 ; Search time 106 Seconds
(without alignments)
4381.659 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTVTSCHLSVAFVPCFL 907

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	1 LGR5_HUMAN	O75473 homo sapien
2	51	5.6	363	2 Q8C8A7	Q8C8A7 mus musculus
3	51	5.6	907	1 LGR5_MOUSE	Q921P4 mus musculus
4	27	3.0	147	2 Q71M96	Q71M96 sus scrofa
5	18	2.0	828	1 LGR6_HUMAN	Q9HXB8 homo sapien
6	18	2.0	915	2 Q6UY15	Q6UY15 homo sapien
7	18	2.0	923	2 Q86V00	Q86V00 homo sapien
8	18	2.0	928	2 Q9BYD7	Q9BYD7 homo sapien
9	15	1.7	878	2 Q8BX59	Q8BX59 mus musculus
10	13	1.4	927	2 Q8N537	Q8N537 homo sapien
11	13	1.4	951	1 LGR4_HUMAN	Q9BXB1 homo sapien
12	13	1.4	951	1 LGR4_RAT	Q9Z2H4 rattus norv
13	11	1.2	134	2 Q80T31	Q80T31 mus musculus
14	11	1.2	162	2 Q80UB8	Q80UB8 mus musculus
15	11	1.2	459	2 Q8R301	Q8R301 mus musculus
16	9	1.0	195	2 Q8UH25	Q8UH25 agrobacteri
17	9	1.0	209	2 Q7D1B7	Q7D1B7 agrobacteri
18	9	1.0	369	2 Q93177	Q93177 caenorhabdi
19	9	1.0	379	2 Q3DG05	Q3DG05 gallus gall
20	9	1.0	424	2 Q9DG06	Q9DG06 gallus gall
21	9	1.0	488	2 Q6PHA3	Q6PHA3 mus musculus
22	9	1.0	589	2 Q9FHL8	Q9FHL8 arabidopsis
23	9	1.0	610	2 Q7GEY4	Q7GEY4 podospora a
24	9	1.0	638	2 Q6Z1I6	Q6Z1I6 oryza sativ
25	9	1.0	653	2 Q8LKV9	Q8LKV9 aegilops ta
26	9	1.0	657	2 Q6ZHK7	Q6ZHK7 oryza sativ
27	9	1.0	693	1 FSHR_CHICK	Q79763 gallus sativ
28	9	1.0	700	2 Q027I4	Q027I4 podospora a
29	9	1.0	869	2 Q32Z53	Q32Z53 arabidopsis
30	9	1.0	871	2 Q7XNV7	Q7XNV7 oryza sativ
31	9	1.0	879	2 Q68C13	Q68C13 oncorhynch

RESULT 1

ID	LGR5_HUMAN	STANDARD;	PRT;	907 AA.
AC	O75473; OSUP75;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Leucine-rich repeat-containing G protein-coupled receptor 5 precursor			
DE	(Orphan G protein-coupled receptor HG38) (G protein-coupled receptor 49).			
GN	Name=GPR49; Synonyms=GPR67, LGR5;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98308104; PubMed=96421114; DOI=10.1006/bbrc.1998.8774;			
RA	McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;			
RT	"Identification and cloning of an orphan G protein-coupled receptor of the-glycoprotein-hormone receptor subfamily.";			
RL	Biochem. Biophys. Res. Commun. 247:266-270(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;			
RA	Hsu S.Y., Liang S.-G., Hsueh A.J.W.;			
RT	"Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";			
RL	Mol. Endocrinol. 12:1830-1845(1998).			
CC	-!- FUNCTION: Orphan receptor. It may be an important receptor for embryonic tissues (by similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal cord, and various region of brain.			
CC	-!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.			
CC	-!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF062006; AAC28019.1; -			
CC	EMBL; AF061444; AAC77911.1; -			
CC	PIR; JE0176; JE0176.			
CC	HSSP; Q9BZR6; IOZN.			
CC	Genew; HGNC:4504; GPR49.			
CC	MIM; 606667; -			

DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR002131; Gphrmn_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF00560; LRR; 16.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00369; LRR_TYP; 8.
 DR SMART; SM00013; LRRNT; 1.
 DR PROSITE; PS00237; G.PROTEIN RECP_F1_1; FALSE_NEG.
 DR PROSITE; PS0262; G.PROTEIN RECP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 907 Leucine-rich repeat-containing G protein-coupled receptor 5.
 FT DOMAIN 1 561 Extracellular (Potential).
 FT TRANSMEM 562 582 1 (Potential).
 FT DOMAIN 583 593 Cytoplasmic (Potential).
 FT TRANSMEM 594 614 2 (Potential).
 FT DOMAIN 615 638 Extracellular (Potential).
 FT TRANSMEM 639 659 3 (Potential).
 FT DOMAIN 660 682 Cytoplasmic (Potential).
 FT TRANSMEM 683 703 4 (Potential).
 FT DOMAIN 704 722 Extracellular (Potential).
 FT TRANSMEM 723 743 5 (Potential).
 FT DOMAIN 744 767 Cytoplasmic (Potential).
 FT TRANSMEM 768 788 6 (Potential).
 FT DOMAIN 789 802 Extracellular (Potential).
 FT TRANSMEM 803 823 7 (Potential).
 FT DOMAIN 824 907 Cytoplasmic (Potential).
 FT REPEAT 64 88 LRR 1.
 FT REPEAT 89 112 LRR 2.
 FT REPEAT 114 136 LRR 3.
 FT REPEAT 137 160 LRR 4.
 FT REPEAT 162 184 LRR 5.
 FT REPEAT 185 208 LRR 6.
 FT REPEAT 209 232 LRR 7.
 FT REPEAT 233 256 LRR 8.
 FT REPEAT 257 279 LRR 9.
 FT REPEAT 281 303 LRR 10.
 FT REPEAT 304 327 LRR 11.
 FT REPEAT 328 350 LRR 12.
 FT REPEAT 351 375 LRR 13.
 FT REPEAT 377 396 LRR 14.
 FT REPEAT 397 420 LRR 15.
 FT REPEAT 422 444 LRR 16.
 FT REPEAT 445 468 LRR 17.
 FT REPEAT 469 492 LRR 18.
 FT REPEAT 493 516 LRR 19.
 FT REPEAT 517 540 LRR 20.
 FT REPEAT 541 564 LRR 21.
 FT CARBOHYD 63 63 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 77 77 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 208 208 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 500 500 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 792 792 N-linked (GlcNAc. . .) (Potential).
 FT CONFLICT 90 90 R -> H (in Ref. 2).
 FT CONFLICT 212 212 L -> W (in Ref. 2).
 SQ SEQUENCE 907 AA; 99997 MW; 822D5C5E6F0D9092 CRC64;

Query Match 100.0%; Score 907; DB 1; Length 907;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
 Db 1 MDTSLGVLISLPVLLQLATGSSPRSGVILRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
 Qy 61 PSNLVSFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120

Db 61 PSNLVSFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
 Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHISYVPPSCFSGLSLRHLWLDNALTEIPVQ 180
 Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHISYVPPSCFSGLSLRHLWLDNALTEIPVQ 180
 Qy 181 AFRSLSALQAWTLALNKIHHIPDYAFGNLSLVVHLHNNRIHISGLKCKCPDGLHSLETLD 240
 Db 181 AFRSLSALQAWTLALNKIHHIPDYAFGNLSLVVHLHNNRIHISGLKCKCPDGLHSLETLD 240
 Qy 241 LNTNNLDEFPPTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIQVGRSA 300
 Db 241 LNTNNLDEFPPTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIQVGRSA 300
 Qy 301 FOHLPELRTLTNGASQITFPPDLTGTANLESLTLTGAQISSSLPQTCVNQLPNIQVLDLS 360
 Db 301 FOHLPELRTLTNGASQITFPPDLTGTANLESLTLTGAQISSSLPQTCVNQLPNIQVLDLS 360
 Qy 361 YNLEDELPSFSVCOKLOKIDLRHNEIYEIKVDTFQQLLSLRSNLNANKIAIHPNAFST 420
 Db 361 YNLEDELPSFSVCOKLOKIDLRHNEIYEIKVDTFQQLLSLRSNLNANKIAIHPNAFST 420
 Qy 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLISSENFPBELKVIEMFYAQCC 480
 Db 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLISSENFPBELKVIEMFYAQCC 480
 Qy 481 AFGVCENAYKI SNQWNGDSSMDLHKDAGMPQAOQDERDLEFLDFEEDLKALHSVQ 540
 Db 481 AFGVCENAYKI SNQWNGDSSMDLHKDAGMPQAOQDERDLEFLDFEEDLKALHSVQ 540
 Qy 541 CSPSPGPKCEHLLDGLIRIGVWTTAVLALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Db 541 CSPSPGPKCEHLLDGLIRIGVWTTAVLALTCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Qy 601 AVNMLTGVSASVAVAGVDAFTFGSPARHAWWENGVCVIGFLSIFASESSVFLTLTAAAL 660
 Db 601 AVNMLTGVSASVAVAGVDAFTFGSPARHAWWENGVCVIGFLSIFASESSVFLTLTAAAL 660
 Qy 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPGE 720
 Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPGE 720
 Qy 721 STMGYVALILLNSLCFLMMTIIATKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
 Db 721 STMGYVALILLNSLCFLMMTIIATKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
 Qy 781 PVAFISFSSLNLTFFISPEVIKFIILLVVPVLPACINPLLYILFNPFKEDLVSLRKQTYV 840
 Db 781 PVAFISFSSLNLTFFISPEVIKFIILLVVPVLPACINPLLYILFNPFKEDLVSLRKQTYV 840
 Qy 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVFTSSITDLPSSVPSPAYPVTESCHLSS 900
 Db 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVFTSSITDLPSSVPSPAYPVTESCHLSS 900
 Qy 901 VAFVPCPL 907
 Db 901 VAFVPCPL 907

RESULT 2
 Q8C8A7 PRELIMINARY; PRT; 363 AA.
 ID Q8C8A7
 AC Q8C8A7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30018C02 product:G protein-coupled receptor 49, full insert sequence.
 DE Name=Gpr49;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

DR Pfam; PF00560; LRR; 16.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PRINTS; PR00019; LRRICHRPT.
DR SMART; SM00369; LRR_TYF; 8.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 907
FT DOMAIN 22 561
FT TRANSMEM 562 582
FT DOMAIN 583 593
FT TRANSMEM 594 614
FT DOMAIN 615 638
FT TRANSMEM 639 659
FT DOMAIN 660 682
FT TRANSMEM 683 703
FT DOMAIN 704 723
FT TRANSMEM 724 744
FT DOMAIN 745 767
FT TRANSMEM 768 788
FT DOMAIN 789 802
FT TRANSMEM 803 823
FT DOMAIN 824 907
FT REPEAT 64 88
FT REPEAT 89 112
FT REPEAT 113 136
FT REPEAT 137 160
FT REPEAT 162 184
FT REPEAT 186 208
FT REPEAT 209 232
FT REPEAT 233 256
FT REPEAT 257 279
FT REPEAT 281 303
FT REPEAT 304 327
FT REPEAT 328 350
FT REPEAT 351 375
FT REPEAT 377 396
FT REPEAT 397 420
FT REPEAT 422 444
FT REPEAT 444 464
FT REPEAT 464 485
FT CARBOHYD 63 63
FT CARBOHYD 77 77
FT CARBOHYD 208 208
FT CARBOHYD 792 792
SQ SEQUENCE 907 AA; 99681 MW; 553167C6C0AAE253 CRC64;

Query Match 5.6%; Score 51; DB 1; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-39;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 203 DYAFGNLSLVVLHNNRIHSLGKCPDGLHSLETLDLNNLNDEPPTAI 253
Db 203 DYAFGNLSLVVLHNNRIHSLGKCPDGLHSLETLDLNNLNDEPPTAI 253

RESULT 4
ID Q71M96 PRELIMINARY; PRT; 147 AA.
AC Q71M96;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE G-protein-coupled receptor 49 (fragment).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Malek M., Ciobanu D.C., Rothschild M.F.; DDBJ databases.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF455789; AAQ04721.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR002131; Gphrmn_receptor.
DR PRINTS; PR00373; GLYCHORMONER.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 147 147
SQ SEQUENCE 147 AA; 15250 MW; 01AE345327EB3586 CRC64;

Query Match 3.0%; Score 27; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 5.6e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 641 GFLSIFASESSVFLTLAALERGFSVK 667
Db 78 GFLSIFASESSVFLTLAALERGFSVK 104

RESULT 5
LGR6 HUMAN
ID LGR6 HUMAN STANDARD; PRT; 828 AA.
AC Q9HBX8; Q96K69;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6.
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549; DOI=10.1210/me.14.8.1257;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duin M., Hsueh A.J.W.;
RA "The three subfamilies of leucine-rich repeat-containing G protein-
RA coupled receptors (LGR): identification of LGR6 and LGR7 and the
RA signaling mechanism for LGR7."
RA Mol. Endocrinol. 14:1257-1271(2000).
RN [2]
RP SEQUENCE OF 406-828 FROM N.A.
RX MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RA "Identification of G protein-coupled receptor genes from the human
RA genome sequence."
RA FEBS Lett. 520:97-101(2002).
RN [3]
RP SEQUENCE OF 431-828 FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
RA Iehida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotsu T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,

```

RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shohata N., Sano S.,
RA Moriya S., Momiya H., Sato N., Takami N., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Takagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Negase T., Nomura N., Kikuchi H., Masuho Y., Yamashita K.,
RA Nakai K., Yada T., Nakamura Y., Ohara G., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF190501; AAG17168.1; -;
DR EMBL; AB083616; BAB89329.1; -;
DR EMBL; AK027377; BAB55071.1; ALT_INIT.
DR HSSP; P25147; 1DOB.
DR Genew; HGNC:19719; LGR6.
DR MIM; 606653; -;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 12.
DR PRINTS; PR00237; GLYCHORMONER.
DR PRINTS; PR00373; GPCR_RHODOPSIN.
DR PRINTS; PR00119; LEURICHRPT.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; FALSE_NEG.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Transmembrane.

FT REPEAT 238 257 LRR 9.
FT REPEAT 258 281 LRR 10.
FT REPEAT 283 305 LRR 11.
FT DISULFID 503 578 By similarity.
FT CARBOHYD 15 15 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).
FT CONFLICT 406 410 CSPTP -> MISPT (in Ref. 2).
FT CONFLICT 628 628 W -> R (in Ref. 3).
FT CONFLICT 824 828 FASHV -> LLHTY (in Ref. 1).
SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AA2D8B4 CRC64;
Query Match 2.0%; Score 18; DB 1; Length 828;
Best Local Similarity 100.0%; Pred. No. 1.1e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;
QY 285 TIHFYDNPQFVGRSAFQ 302
|||||
DB 146 TIHFYDNPQFVGRSAFQ 163
RESULT 6
Q6UY15 PRELIMINARY; PRT; 915 AA.
AC Q6UY15
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Gonadotropin receptor.
GN ORFNames=UNQ6427;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22897296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY358119; AAQ88486.1; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00119; LEURICHRPT.
DR SMART; SM00365; LRR_SD22; 5.
DR SMART; SM00369; LRR_TYP; 14.
KW Receptor.
SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DB8555F4 CRC64;
Query Match 2.0%; Score 18; DB 2; Length 915;
Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;
QY 285 TIHFYDNPQFVGRSAFQ 302

```

|||||
233 TIHFYDNPQFVGRSAFQ 250

Db
RESULT 7
Q86VU0
ID Q86VU0 PRELIMINARY; PRT; 923 AA.
AC Q86VU0,
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL EMBL; BC047905; AAH47905.1; -.
DR HSSP; Q9BZR6; 1P8T.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Ghrmn_receptor.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON_TER 1
SQ SEQUENCE 923 AA; 10047 MW; 541D6746DAB06813 CRC64;

Query Match 2.0%; Score 18; DB 2; Length 923;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 TIHFYDNPQFVGRSAFQ 302
|||||
Db 241 TIHFYDNPQFVGRSAFQ 258

RESULT 8
Q86XS9
ID Q86XS9 PRELIMINARY; PRT; 878 AA.
AC Q86XS9,
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DE clone:A930003A08 product:G PROTEIN-COUPLED RECEPTOR LGRA homolog
DE (Fragment).
DE (Fragment).
GN Name=Gpr48;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.

```


RC STRAIN=C57BL/6J; TISSUE=Retina;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:11617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishii K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:11757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK044357; BAC31882.1; -
DR HSSP; P25147; 1D08.
DR MGD; MGI:1891468; Gpr48.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gpr48m_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_TYP.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_1; 14.
DR PRINTS; PR00373; GYCHORMONER.
DR PRINTS; PR00237; GPCRHOOPS.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Receptor.
FT NON_TER.
FT SEQUENCE 878 AA; 96877 MW; 68E64B5E5DEA11B37 CRC64;
Query Match 1.7%; Score 15; DB 2; Length 878;
Best Local Similarity 100.0%; Pred. No. 8.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 LETLDLNNLDEFP 250

Db 154 LETLDLNNLDEFP 168
RESULT 10
Q8N537
ID Q8N537 PRELIMINARY; PRT; 927 AA.
AC Q8N537;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GPR48 protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC033039; AAH33039.1; -
DR HSSP; P25147; 1D08.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gpr48m_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_TYP.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00560; LRR_1; 14.
DR PRINTS; PR00373; GYCHORMONER.
DR PRINTS; PR00237; GPCRHOOPS.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 927 AA; 101676 MW; C7E2F0C40E584CE8 CRC64;
Query Match 1.4%; Score 13; DB 2; Length 927;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 208 NLSSLVVLHNN 220

```

Db      175 NLSSLVVLHNN 187
|||||
RESULT 11
LGR4 HUMAN
ID LGR4 HUMAN STANDARD; PRT; 951 AA.
AC Q9XB1; Q9YD1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
DE (G protein-coupled receptor 48).
GN Name=GPR48; Synonyms=LGR4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=21294803; PubMed=11401528; DOI=10.1006/bbrc.2001.4625;
RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT "Molecular characterization of a novel glycoprotein hormone G-protein-
coupled receptor.";
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
placenta, ovary, testis and adrenal. Expressed also in spinal
cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
and spleen.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF346711; AAK31153.1; -
EMBL; AF346709; AAK31153.1; JOINED.
EMBL; AF346710; AAK31153.1; JOINED.
EMBL; AF257182; AAF68989.1; -
DR HSP; O9BZR6; 102N.
DR Genew; HGNC:13299; GPR48.
DR MIM; 606666; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 15.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GLYCHORMONER.
DR PRINTS; PR00373; GPCR_RHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 4.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 951 Leucine-rich repeat-containing G protein-
coupled receptor 4.
FT DOMAIN 25 544 Extracellular (Potential).
FT TRANSMEM 545 565 1 (Potential).
FT DOMAIN 566 575 Cytoplasmic (Potential).

```

```

FT TRANSMEM 576 596 2 (Potential).
FT DOMAIN 597 620 Extracellular (Potential).
FT TRANSMEM 621 641 3 (Potential).
FT DOMAIN 642 661 Cytoplasmic (Potential).
FT TRANSMEM 662 682 4 (Potential).
FT DOMAIN 683 703 Extracellular (Potential).
FT TRANSMEM 704 724 5 (Potential).
FT DOMAIN 725 756 Cytoplasmic (Potential).
FT TRANSMEM 757 777 6 (Potential).
FT DOMAIN 778 783 Extracellular (Potential).
FT TRANSMEM 784 804 7 (Potential).
FT DOMAIN 805 951 Cytoplasmic (Potential).
FT REPEAT 55 79 LRR 1.
FT REPEAT 81 103 LRR 2.
FT REPEAT 104 127 LRR 3.
FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 201 223 LRR 7.
FT REPEAT 224 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 368 387 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 By similarity.
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
FT CONFLICT 292 292 F -> S (in Ref. 1; AAF68989).
FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Query Match 1.4%; Score 13; DB 1; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 NLSSLVVLHNN 220
DB 199 NLSSLVVLHNN 211
|||||
RESULT 12
LGR4 RAT
ID LGR4 RAT STANDARD; PRT; 951 AA.
AC Q92ZH4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN Name=Gpr48; Synonyms=Lgr4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
thyrotropin receptors with extracellular leucine-rich repeats and a G
protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.

```



```
FT  NON TER      1      1
SQ  NON TER     162     162
SQ  SEQUENCE   162 AA;  17608 MW;  C593128C551FB824 CRC64;

Query Match      1.2%; Score 11; DB 2; Length 162;
Best Local Similarity 100.0%; Pred.No. 0.15;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  810 PLPACLNPLLY 820
Db  70 PLPACLNPLLY 80
|||||
70 PLPACLNPLLY 80

RESULT 15
Q8R301
ID  Q8R301      PRELIMINARY;      PRT;   459 AA.
AC  Q8R301;
DT  01-JUN-2002 (TReMBLrel. 21, Created)
DT  01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT  01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE  Lgr6 protein (Fragment).
GN  Name=Lgr6;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Czech II; TISSUE=Mammary tumor;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Czech II; TISSUE=Mammary tumor;
RA  Strausberg R.;
RL  Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC026896; AAH26896.1; -.
DR  MGD; MGI:2441805; Lgr6.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR  InterPro; IPR000276; GPCR_Rhodopsn.
DR  InterPro; IPR002131; Gqhrma_receptor.
DR  Pfam; PF00001; 7tm1; 1.
DR  PRINTS; PR00373; GLYCHORMONER.
DR  PRINTS; PR00237; GPCRHHODOPSN.
FT  NON TER      1
SQ  SEQUENCE   459 AA;  47889 MW;  F0100BF073E81762 CRC64;

Query Match      1.2%; Score 11; DB 2; Length 459;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  810 PLPACLNPLLY 820
```

Db 309 PLPACLNPLLY 319
|||||

Search completed: July 12, 2005, 07:53:08
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:48:53 ; Search time 25 Seconds
(without alignments)
2708.268 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLRGVLLSLPVLLQLAT.....PAYPTESCHLSVAVPCL 907

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2.6/prodata/1/iaa/5A COMB.pep:*
 - 2: /cgn2.6/prodata/1/iaa/5B COMB.pep:*
 - 3: /cgn2.6/prodata/1/iaa/6A COMB.pep:*
 - 4: /cgn2.6/prodata/1/iaa/6B COMB.pep:*
 - 5: /cgn2.6/prodata/1/iaa/PCTUS COMB.pep:*
 - 6: /cgn2.6/prodata/1/iaa/backfileseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	4	US-09-170-496D-264
2	806	88.9	907	4	US-09-170-496D-278
3	9	1.0	644	2	US-08-866-757-2
4	9	1.0	644	3	US-09-153-593-2
5	9	1.0	1133	4	US-10-101-464A-809
6	8	0.9	99	4	US-09-482-273-258
7	8	0.9	207	4	US-09-482-273-167
8	8	0.9	232	4	US-09-543-681A-5153
9	8	0.9	259	4	US-09-482-273-260
10	8	0.9	386	4	US-09-489-039A-10334
11	8	0.9	448	4	US-09-361-443-4
12	8	0.9	467	4	US-09-361-443-2
13	8	0.9	469	4	US-09-198-452A-793
14	8	0.9	469	4	US-09-438-185A-745
15	8	0.9	728	4	US-09-949-016-10391
16	7	0.8	28	3	US-09-099-041A-19
17	7	0.8	28	3	US-09-245-281-19
18	7	0.8	28	4	US-09-207-359B-19
19	7	0.8	28	4	US-09-340-620A-19
20	7	0.8	28	4	US-09-865-364-19
21	7	0.8	36	4	US-08-216-592A-22
22	7	0.8	67	4	US-09-248-796A-21331
23	7	0.8	69	4	US-09-910-009A-158
24	7	0.8	88	4	US-09-248-796A-23648
25	7	0.8	95	4	US-09-328-352-5647
26	7	0.8	96	4	US-09-513-999C-4592
27	7	0.8	97	4	US-09-513-999C-7456

28	7	0.8	113	4	US-08-816-772-2	Sequence 2, Appli
29	7	0.8	115	4	US-09-640-211A-2235	Sequence 2235, Ap
30	7	0.8	127	3	US-09-134-001C-4041	Sequence 4041, Ap
31	7	0.8	135	4	US-10-101-464A-563	Sequence 563, App
32	7	0.8	137	4	US-09-489-039A-11223	Sequence 11223, A
33	7	0.8	149	4	US-09-902-540-11269	Sequence 11269, A
34	7	0.8	154	4	US-09-540-236-2694	Sequence 2694, Ap
35	7	0.8	156	4	US-09-252-991A-23000	Sequence 23000, A
36	7	0.8	157	4	US-09-270-767-40195	Sequence 40195, A
37	7	0.8	157	4	US-09-270-767-55411	Sequence 55411, A
38	7	0.8	161	4	US-09-248-796A-15595	Sequence 15595, A
39	7	0.8	161	4	US-10-101-464A-570	Sequence 570, App
40	7	0.8	166	4	US-09-252-991A-21079	Sequence 21079, A
41	7	0.8	166	4	US-09-889-914B-8	Sequence 8, Appli
42	7	0.8	168	4	US-09-270-767-57047	Sequence 57047, A
43	7	0.8	170	3	US-08-875-573-6	Sequence 6, Appli
44	7	0.8	176	4	US-09-134-000C-4892	Sequence 4892, Ap
45	7	0.8	179	4	US-09-252-991A-24661	Sequence 24661, A

ALIGNMENTS

RESULT 1

US-09-170-496D-264
; Sequence 264, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 264
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-264

Query Match 100.0%; Score 907; DB 4; Length 907;

Best Local Similarity	100.0%;	Pred. No. 0;
Matches 907;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
QY	61	PNLSVFTSYLDLNMNITSQLPNPLPSLRFLEELRAGNALTYPKGFTGLYSKVLV 120
Db	61	PNLSVFTSYLDLNMNITSQLPNPLPSLRFLEELRAGNALTYPKGFTGLYSKVLV 120
QY	121	LQNNLRHVPTALQNLRSLSQSLRDANHISVVPSCFGLSHSLRHLWDDNALTIPVQ 180
Db	121	LQNNLRHVPTALQNLRSLSQSLRDANHISVVPSCFGLSHSLRHLWDDNALTIPVQ 180
QY	181	AFRSLAQAMTALANKIHHPDYAFGNLSLVTLHLHNNRHSLSGKCFDGLHSLSETLD 240
Db	181	AFRSLAQAMTALANKIHHPDYAFGNLSLVTLHLHNNRHSLSGKCFDGLHSLSETLD 240
QY	241	LYNNMLDEPPTAIRLSNLKELGFTSNNSIRSIPEKAFVGNPSLTIHFVDNPIQFVGSA 300
Db	241	LYNNMLDEPPTAIRLSNLKELGFTSNNSIRSIPEKAFVGNPSLTIHFVDNPIQFVGSA 300
QY	301	FOHLPRLTLTNGASQITTEFPDLTGTLNLSLTITGAQISLIPQVNCNQLPNQLVLDLS 360
Db	301	FOHLPRLTLTNGASQITTEFPDLTGTLNLSLTITGAQISLIPQVNCNQLPNQLVLDLS 360
QY	361	YNLLEDLFSVSCQKQLKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHPNAFST 420

Db 361 YNLEDELFVSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIHHNPFST 420
Qy 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKLGNHALQSLISSENPELKVIEPMPYAYQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKLGNHALQSLISSENPELKVIEPMPYAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSVAVLGADFTFGSFARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSVAVLGADFTFGSFARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATWAAVPLLGGSYKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATWAAVPLLGGSYKYGASPLCLPFGEP 720
Qy 721 STMGYMWALILLNSLCFLMWTIATYTKLYCNLDKGDLENIDCSMWKHIALLFTNCILNC 780
Db 721 STMGYMWALILLNSLCFLMWTIATYTKLYCNLDKGDLENIDCSMWKHIALLFTNCILNC 780
Qy 781 PVAFLSFSLSLINLTFTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db 781 PVAFLSFSLSLINLTFTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 2
US-09-170-496D-278
; Sequence 278, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-278

Query Match 88.9%; Score 806; DB 4; Length 907;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MDTSRIGVLLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSL 60
Db 1 MDTSRIGVLLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSL 60
Qy 61 PSNLSVFTSYLDLSMNINSQLLPNPLSRFLFEEELRLAGNALTYIPKGFTGLYSKVL 120
Db 61 PSNLSVFTSYLDLSMNINSQLLPNPLSRFLFEEELRLAGNALTYIPKGFTGLYSKVL 120

Qy 121 LONNOLRHVTEALQNLRSLOSLRDANHI SYVPPSCFSGHLSRLHLWLDNALTEIPVQ 180
Db 121 LONNOLRHVTEALQNLRSLOSLRDANHI SYVPPSCFSGHLSRLHLWLDNALTEIPVQ 180
Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSIGKCKFCGLHSLFETLD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSIGKCKFCGLHSLFETLD 240
Qy 241 LNYNNLDFPPTAIRTLNKLGLGPHSNNIRSIPEKAFVGNFSLTIHFDNPIQVGRSA 300
Db 241 LNYNNLDFPPTAIRTLNKLGLGPHSNNIRSIPEKAFVGNFSLTIHFDNPIQVGRSA 300
Qy 301 FOHLPELRTLTNGASQITTEPDLTGTAANLSLTGTGAQISSLPOTVCNQLPNLOVLDLS 360
Db 301 FOHLPELRTLTNGASQITTEPDLTGTAANLSLTGTGAQISSLPOTVCNQLPNLOVLDLS 360
Qy 361 YNLEDELFVSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIHHNPFST 420
Db 361 YNLEDELFVSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIHHNPFST 420
Qy 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKLGNHALQSLISSENPELKVIEPMPYAYQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKLGNHALQSLISSENPELKVIEPMPYAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSVAVLGADFTFGSFARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSVAVLGADFTFGSFARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATWAAVPLLGGSYKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATWAAVPLLGGSYKYGASPLCLPFGEP 720
Qy 721 STMGYMWALILLNSLCFLMWTIATYTKLYCNLDKGDLENIDCSMWKHIALLFTNCILNC 780
Db 721 STMGYMWALILLNSLCFLMWTIATYTKLYCNLDKGDLENIDCSMWKHIALLFTNCILNC 780
Qy 781 PVAFLSFSLSLINLTFTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db 781 PVAFLSFSLSLINLTFTISPEVIFKILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPECL 907
Db 901 VAFVPECL 907
RESULT 3
US-08-866-757-2
; Sequence 2, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGMA, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA


```
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-866-757-2

Query Match 1.0%; Score 9; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 PYAYQCCAF 482
Db 158 PYAYQCCAF 166

RESULT 4
US-09-153-593-2
; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSA, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CNA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-153-593-2

Query Match 1.0%; Score 9; DB 3; Length 644;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 PYAYQCCAF 482
Db 158 PYAYQCCAF 166

RESULT 5
US-10-101-464A-809
; Sequence 809, Application US/10101464A

; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 809
; LENGTH: 1133
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
; US-10-101-464A-809

Query Match 1.0%; Score 9; DB 4; Length 1133;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 LQVLDLSYN 362
Db 175 LQVLDLSYN 183

RESULT 6
US-09-482-273-258
; Sequence 258, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-482-273-258

Query Match 0.9%; Score 8; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 YTKLYCNL 751
Db 33 YTKLYCNL 40

RESULT 7
US-09-482-273-167
; Sequence 167, Application US/09482273
```

; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-167

Query Match 0.9%; Score 8; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 PLPACLPN 817
|||||||
DB 47. PLPACLPN 54

RESULT 8

US-09-543-681A-5153
; Sequence 5153, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5153
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5153

Query Match 0.9%; Score 8; DB 4; Length 232;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 FTSYLDLS 74
|||||||
DB 220 FTSYLDLS 227

RESULT 9

US-09-482-273-260
; Sequence 260, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-260

Query Match 0.9%; Score 8; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 YTKLYCNL 751
|||||||
DB 33 YTKLYCNL 40

RESULT 10

US-09-489-039A-10334
; Sequence 10334, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10334
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10334

Query Match 0.9%; Score 8; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 DLVSLRKQ 837
|||||||
DB 352 DLVSLRKQ 359

RESULT 11

US-09-361-443-4
; Sequence 4, Application US/09361443
; Patent No. 6660275
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Omen, Raymond P
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-005 (AV-5)
; CURRENT APPLICATION NUMBER: US/09/361,443
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/094,195
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-361-443-4

```
Query Match          0.9%; Score 8; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 786 SFSSLINL 793
DB 290 SFSSLINL 297

RESULT 12
US-09-361-443-2
; Sequence 2, Application US/09361443
; Patent No. 6660275
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Oomen, Raymond P
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 19721-005 (AV-5)
; CURRENT APPLICATION NUMBER: US/09/361,443
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/094,195
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-361-443-2

Query Match          0.9%; Score 8; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 786 SFSSLINL 793
DB 292 SFSSLINL 299

RESULT 13
US-09-198-452A-793
; Sequence 793, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 793
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-793

Query Match          0.9%; Score 8; DB 4; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 786 SFSSLINL 793
DB 294 SFSSLINL 301

RESULT 14
US-09-438-185A-745
; Sequence 745, Application US/09438185A
; Patent No. 6822071
```

```
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-00041105
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 745
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0743
US-09-438-185A-745

Query Match          0.9%; Score 8; DB 4; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 786 SFSSLINL 793
DB 294 SFSSLINL 301

RESULT 15
US-09-949-016-10391
; Sequence 10391, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10391
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10391

Query Match          0.9%; Score 8; DB 4; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 650 SSVFLLTL 657
DB 580 SSVFLLTL 587

Search completed: July 12, 2005, 07:54:11
Job time : 28 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: July 12, 2005, 07:34:30 ; Search time 31 Seconds
(without alignments)
2815.116 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGVLSPVLLQLAT.....PAYPTESCHLSSVAFVCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	2 JG0176	orphan G protein-c
2	4027	85.6	907	2 JG0193	G protein-coupled
3	650	13.8	695	2 I45896	follicle stimulat
4	648.5	13.8	695	1 JN0898	follicle stimulat
5	642.5	13.7	694	2 JC2237	follicle stimulat
6	642	13.7	695	2 JC1493	follicle stimulat
7	636	13.5	692	2 A34548	follicle stimulat
8	632	13.4	694	2 JC4301	follicle stimulat
9	625	13.3	696	2 JC7361	follicle stimulat
10	621.5	13.2	695	1 QRHUFT	follicle stimulat
11	600	12.8	700	2 A42395	lutropin receptor
12	594	12.6	764	1 ORHURH	thyrotropin recept
13	591	12.6	764	2 A40077	thyrotropin recept
14	588.5	12.5	700	2 I77463	lutropin-choriogon
15	588.5	12.5	700	2 A49744	thyrotropin recept
16	581	12.4	764	2 I48882	thyrotropin recept
17	578	12.3	793	2 JC7390	thyroid stimulat
18	572	12.2	696	2 A41344	lutropin-choriogon
19	567.5	12.1	764	2 JC5643	thyroid stimulat
20	566.5	12.0	764	2 A35956	thyrotropin recept
21	560	11.9	699	1 ORHUFT	lutropin-choriogon
22	536.5	11.4	814	2 JC7389	thyroid stimulat
23	508.5	10.8	889	2 T20123	hypothetical prote
24	490	10.4	925	2 JC2033	G protein-coupled
25	426.5	9.1	603	2 JC1282	insulin-like growt
26	421	9.0	605	2 JC5239	insulin-like growt
27	418.5	8.9	622	2 JC7973	synleucin - human
28	406.5	8.6	603	2 JC6128	insulin-like growt
29	406.5	8.6	1389	2 TJ3852	gene wheeler prote

30	403	8.6	1523	2 T13953	MEGF5 protein - ra
31	402	8.5	605	2 A41915	insulin-like growt
32	400.5	8.5	1385	2 T13887	tlr protein - frui
33	398	8.5	1091	2 A58532	glial cell membran
34	391.5	8.3	1469	2 B36665	slit protein 2 pre
35	391.5	8.3	1480	2 A36665	slit protein 1 pre
36	385.5	8.2	1531	2 T42218	slit-1 protein hom
37	382.5	8.1	1066	2 T15864	hypothetical prote
38	381.5	8.1	1115	2 S40241	G protein-coupled
39	371	7.9	1119	2 AD1822	leucine-rich-repea
40	370	7.9	559	2 T42998	ras-binding protei
41	370	7.9	560	2 A60164	platelet membrane
42	364.5	7.8	572	2 T30947	hypothetical prote
43	360.5	7.7	1134	1 A29444	chaoptin precursor
44	356.5	7.6	536	2 A34901	lysine carboxypept
45	354.5	7.5	707	2 JC7763	neuronal leucine-r

ALIGNMENTS

RESULT 1

JE0176 orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man)

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004

C:Accession: JE0176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A:Title: Identification and cloning of an orphan G protein-coupled receptor of the glyco-

A:Reference number: JE0176; MUID:98308104; PMID:9642114

A:Accession: JE0176

A:Molecule type: mRNA

A:Residues: 1-907 <MCD>

A:Cross-references: UNIPROT:O75473; GB:AF062006; NID:G3366801; PIDN:AAC28019.1; PID:G33

C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C:Genetics:

A:Gene: HG38

A:Map position: 12Q22-23

F:1-21/Domain: signal sequence #status predicted <SIG>

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:639-660/Domain: transmembrane #status predicted <TM3>

F:681-701/Domain: transmembrane #status predicted <TM4>

F:725-744/Domain: transmembrane #status predicted <TM5>

F:768-791/Domain: transmembrane #status predicted <TM6>

F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 100.0%; Score 4702; DB 2; Length 907;

Best Local Similarity 100.0%; Pred. No. 4.6e-297;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60
DB	1	MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60
QY	61	PSNLISVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLVM	120
DB	61	PSNLISVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLVM	120
QY	121	LQNNQLRHVPTEALQNLRLDANHI SVVPPSCFSGLSRLHRLWDDNALTEIPVQ	180
DB	121	LQNNQLRHVPTEALQNLRLDANHI SVVPPSCFSGLSRLHRLWDDNALTEIPVQ	180
QY	181	AFRSLSALQMTALANKIHHIPDYAFGNLSLVLLHNNRHSLSGKKCFDGLHSLTLD	240
DB	181	AFRSLSALQMTALANKIHHIPDYAFGNLSLVLLHNNRHSLSGKKCFDGLHSLTLD	240
QY	241	LYNNNLDEFPTAIRTLNLKELGPHSNIRSIPEKAFVGNPNLSLI THFYDNP IQVGRSA	300
DB	241	LYNNNLDEFPTAIRTLNLKELGPHSNIRSIPEKAFVGNPNLSLI THFYDNP IQVGRSA	300
QY	301	FQHLPELRTLUNGASQITEFPDLTGTANLESLLTGAQISSLPQTCVNCNQLPNLQVLDLS	360

```
Db 301 FOHLELRTLTNGASQITTEPDLTGANLESRLTGAISSLQTCVQCNLPNLQVLDLS 360
Qy 361 YNLEDLPSFVCQKQKIDLRHNEIYEIKVDITFQQLSLSLSLNLANWKIAIHPNAPST 420
Db 361 YNLEDLPSFVCQKQKIDLRHNEIYEIKVDITFQQLSLSLSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Db 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Qy 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSVAVLAVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Db 601 AVNMLTGVSVAVLAVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Qy 661 ERGSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGVMVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGVMVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSLLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHFKEDLVSLRKQTYV 840
Db 781 PVAFLSPSSLLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHFKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDVVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVVTTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 2
JG0193
G protein-coupled receptor FEX - mouse
C;Species: Mus musculus (house mouse)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C;Accession: JG0193
R;Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A;Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin
A;Reference number: JG0193; MUID:99121227; PMID:9920770
A;Accession: JG0193
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-907 <HER>

Query Match 85.6%; Score 4027; DB 2; Length 907;
Best Local Similarity 85.3%; Pred No. 2,8e-253;
Matches 774; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQIATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60
Db 1 MDTSCVHMLLSLALLQLVAAGSPGDAIPRGCPSHCHCELDGRMLLRVDCSDLGLSEL 60
Qy 61 PSNLSVFTSYDLSNMNITSOILLPNPLPSIRFLEELRLAGNALTYPKCAFTGLSKVLIM 120
Db 61 PSNLSVFTSYDLSNMNITSOILLPNPLPSIRFLEELRLAGNALTYPKCAFTGLSKVLIM 120
Qy 121 LQNNOLRHPVEALQNLRLSLQSLRDANHISYVPPSCFSGHLHLWLDNALTDEIPVQ 180
Db 121 LQNNOLRHPVEALQNLRLSLQSLRDANHISYVPPSCFSGHLHLWLDNALTDEIPVQ 180
```

```
Db 121 LQNNOLRHPVEALQNLRLSLQSLRDANHISYVPPSCFSGHLHLWLDNALTDEIPVQ 180
Qy 181 AFRSLSALQAWTLALNKIHHIPDYAFGNLSLAVLHLHNNRIHSGIKKCFGLHSLETLD 240
Db 181 AFRSLSALQAWTLALNKIHHIADYAFGNLSLAVLHLHNNRIHSGIKKCFGLHSLETLD 240
Qy 241 LNYNNLDEFFTAIRTLNKLKELGPHSNNIRSIPEKAFVGNPSLTIITHFYDNPFIQFVGSA 300
Db 241 LNYNNLDEFFTAIRTLNKLKELGPHSNNIRSIPEKAFVGNPSLTIITHFYDNPFIQFVGSA 300
Qy 301 FOHLELRTLTNGASQITTEPDLTGANLESRLTGAISSLQTCVQCNLPNLQVLDLS 360
Db 301 FOHLELRTLTNGASQITTEPDLTGANLESRLTGAISSLQTCVQCNLPNLQVLDLS 360
Qy 361 YNLEDLPSFVCQKQKIDLRHNEIYEIKVDITFQQLSLSLSLNLANWKIAIHPNAPST 420
Db 361 YNLEDLPSFVCQKQKIDLRHNEIYEIKVDITFQQLSLSLSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Db 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Qy 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSVAVLAVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Db 601 AVNMLTGVSVAVLAVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLITLAAL 660
Qy 661 ERGSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGVMVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGVMVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSLLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHFKEDLVSLRKQTYV 840
Db 781 PVAFLSPSSLLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHFKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDVVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVVTTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 3
I45896
follicle stimulating hormone receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C;Accession: I45896
R;Houde, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.
Mol. Reprod. Dev. 39, 127-135, 1994
A;Title: Structure of the bovine follicle-stimulating hormone receptor complementary DNA
A;Reference number: I45896; MUID:95127199; PMID:7826612
A;Accession: I45896
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-895 <HOU>
C;Cross-references: UNIPROT:P35376; GB:I22319; NID:g404671; PIDN:AAC37324.1; PID:g404672
C;Genetics:
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
```



```
QY 673 ETKAPSSKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEPSTWGVVALILL 732
Db 477 QCKQVRRHAASIMLYGWIFAFTVALFPIFGISSYMKVSIICLPMDIDSPLSQLYVVSLLVL 536
QY 733 NSLCFLMWTIAVTKLYCNL-DKGDLENTWDCSMVKHIALLLFTNCILNCPVAFLSFSLI 791
Db 537 NVLAFVILGCVIHYLYVRPNINSSSDYKIAKEMALFTDFLCWVPISFFAISASL 596
QY 792 NLTFISPEVIRKIPILLVVPPLPACMLPLLYLPNPHFKEDLDVLSKQTYVWTRSKHPSLMS 851
Db 597 KVPLITVSKILLVLPVINSANPFLYIFTKFNFRDVFIL----- 639
QY 852 INSDVEKQSDSTQALVTFSS-SITVDLPSSVSPSPAYP-VTESCHLSSVAFVP 905
Db 640 -----LSKFGCYEMQAQTYRTENLSTAHNIHPRNGHCPAPRITNS-----SSTLIP 687

RESULT 9
JC7361
folliotropin receptor precursor - new
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Cynops pyrrhogaster (new)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: JC7361
R:Nakayama, Y.; Yamamoto, T.; Oba, Y.; Nagahama, Y.; Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A:Title: Molecular cloning, functional characterization, and gene expression of a follitropin receptor from the Chinese giant salamander, Cynops pyrrhogaster
A:Reference number: JC7361
A:Contents: Testis
A:Accession: JC7361
A:Molecule type: mRNA
A:Residues: 1-696 <NA>
A:Cross-references: UNIPROT:Q9DGF5; DBJ:AB005587
C:Comment: This protein, containing seven transmembrane domains and a large glycosylated in and thyroid stimulating hormone. This receptor has a common signal transduction pathway and thyroid stimulating hormone.
C:Genetics:
A:Gene: fsh-r
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; transmembrane
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-696/Product: follicle-stimulating hormone receptor #status predicted <MAT>
F:18-359/Domain: extracellular #status predicted <EXT>
F:370-389/Domain: transmembrane #status predicted <TM1>
F:402-424/Domain: transmembrane #status predicted <TM2>
F:447-468/Domain: transmembrane #status predicted <TM3>
F:489-511/Domain: transmembrane #status predicted <TM4>
F:532-553/Domain: transmembrane #status predicted <TM5>
F:577-600/Domain: transmembrane #status predicted <TM6>
F:612-633/Domain: transmembrane #status predicted <TM7>
F:46,190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:445-520/Disulfide bonds: #status predicted

Query Match 13.3%; Score 625; DB 2; Length 696;
Best Local Similarity 22.0%; Pred. No. 1e-32;
Matches 200; Conservative 140; Mismatches 313; Indels 256; Gaps 23;

QY 10 LSLPVLQLATGGSPRGVLLRGCPTRCHCEPDGRMLLRV-DCSDLGLSLPNSLVFT 68
Db 1 MSLAILCLLLAVGS-----FGCHPVCR-----LNRVPTCQESHVQIPR----- 41
QY 69 SYLDSMNINISQLPNSLRFDELRLAGNALYIPKGAFTGLYSIK-VLMQLNNQLR 127
Db 42 ---DIPRNS-----TELRFVLTKVTIPKAAFGSPEDVENIEISQNDVLK 83
QY 128 HVPTEALQNLRSLOSLRLD-ANHSISYVPSFCFSGHLSRLHLDNDALETPVQAFRLS 186
Db 84 TIEANVFSHLPKLRIRIEKANNLYIDPFAQNLPSLYKLLISNTGTLQVLP----- 135
QY 187 ALQAMTLALNKHIPDPVAFGNLSLVVHLHNN-RIHSLGKCKCPDGLSHSLTLDLNNY 245
Db 136 -----AVSKIRSFHS-----VLLDVQDNINIRHIGKNSFAGLSS-ESTIRLN- 177
QY 246 LDEPPTAIRTLNLKELGFHSHNNIRSIPEKAFVGNPNSLITTHFYDNPNIQFVGRSAFQHL 305
```

```
Db 178 -----KNGIEEQNHAFNGT-----HUN 195
QY 306 ELRITLTNGASQITEFFDLTGTANLESULTGAQISSLPQTVCNQLPNLQVLDLSYNLLE 365
Db 196 E-----LNLSDN----- 202
QY 366 DLPFSVCQKLOKIDLRNEIYKVDTFQOLLSLRSLNLAWNKAITHPNAPFSLPSLI 425
Db 203 -----QRLKLP-----DQVFOGATGVP 220
QY 426 KLDLSSNLLSFPITGLHGLTHLKTGNHALQSLSSSENEPELKVIEMPYAQQCAFG-- 483
Db 221 ILDISRTRIHFLPNNGIENIKKFRARFNYLLKKLPPELKPALIEANLITYPSHCAPANR 280
QY 484 -----VCENAYKISNOWNKGDNSMDDLHKDKAGMFOAQDRDLDFLLDFFED-- 532
Db 281 ERKSEMHPICNKSF-----GKHDSAEKPEDKTLRRFSNEDYLSVSGFSYSLVNGD 332
QY 533 -----LKALHSVQCSPPSPGPKPCHEHLLDGLMIRIGVMTIAVLATCNALVTSTVPR 584
Db 333 BFNVDYILCNEVHDVICFPKPDAPNCPEDIMGDNTRLRLIWLISILAITGNITVLVILIS 392
QY 585 SPLISPIKLIGVIAAVNMLTVSSAVLAGVDAFTFGSFARHGAWENGVCCHVIGFLS 644
Db 393 SOYFTVPRFLMCLNAPADLCWGIYLLLASVDIKTSQYNNHADMOTGSCAAAGFT 452
QY 645 IFASESSVFLTLAALERGFVKTSAPETKAPSSSLKVIILLCALLALTMAAVPLLAGS 704
Db 453 VFASLSVYTLVTITIERWHITVAMQLDKRFRHATAIMVVGWIFATVAILPIFGVS 512
QY 705 KYGASPLCLPFPGEPSMTGMVVALIILNSLCFLMWTIAVTKLYCNL-DKGDLENWDCS 763
Db 513 SYTKVSIICLPMDIESLSQSIVIVLNLVCAFLIICACIYIGYLTVRPNVSVNSDTK 572
QY 764 MVKHALLFTNCILNCPVAFLSFSSLINLFIPEVTKFILLVVPPLPACMLPLLYLIF 823
Db 573 IAKEMALIFTDFLCMAPIPSFAISAKIPLITVSKSKILLVLPYNSCANPLFYAIF 632
QY 824 NPHEKEDLVLSRKQTYVWTRSKHPSLMSINDVVEKQSDSTQALVTFSSSI--TYDLP 881
Db 633 TKTFRRDFIL-----MSKFGCEVQAQNCR-----TETSSSLHSHNRN 672
QY 882 PSSVSPSPAY 890
Db 673 GHYVPAPKY 681
```

RESULT 10

QKHUFT

folliotropin receptor precursor - human

N:Alternate names: follicle stimulating hormone receptor (FSHR)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1991 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004

C:Accession: 157661; I56448; PC1147; S30560; I57672; JN0122

R:Gromoll, J.; Dankbar, B.; Gudermann, T.

Mol. Cell. Endocrinol. 102, 93-102, 1994

A:Title: Characterization of the 5' flanking region of the human follicle-stimulating ho

A:Reference number: 157661; MUID:95011044; PMID:7926278

A:Accession: 157661

A:Molecule type: DNA

A:Molecule type: DNA

A:Residues: 1-51 <GRO>

A:Cross-references: UNIPROT:P23945; GB:S73199; NID:g685036; PIDN:AAB32071.1; PID:g685037

R:Gromoll, J.; Ried, T.; Holtgreve-Grez, H.; Nieschlag, E.; Gudermann, T.

J. Mol. Endocrinol. 12, 265-271, 1994

A:Title: Localization of the human FSH receptor to chromosome 2 p21 using a genomic prob

A:Reference number: 156448; MUID:95000244; PMID:7916967

A:Accession: 156448

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 286-695 <GR2>

A;Cross-references: GB:S73526; NID:G688069; PIDN:AA32225.1; PID:G688070
R;Gromoll, J.; Gudermann, T.; Nieschlag, E.
Biochem. Biophys. Res. Commun. 186, 1077-1083, 1992
A;Title: Molecular cloning of a truncated isoform of the human follicle stimulating hormone
A;Reference number: PC1147; MUID:93075197; PMID:1359889
A;Accession: PC1147
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-223,286-294, 'P', 296-342 <GR3>
A;Cross-references: EMBL:X68044; NID:G31473; PIDN:CAA48179.1; PID:G31474
A;Experimental source: testis
R;Gromoll, J.
submitted to the EMBL Data Library, August 1992
A;Reference number: S30560
A;Accession: S30560
A;Molecule type: mRNA
A;Residues: 1-12, 'R', 14-223,286-294, 'P', 296-342 <GR4>
A;Cross-references: EMBL:X68044; NID:G31473; PIDN:CAA48179.1; PID:G31474
R;Kelton, C.A.; Cheng, S.V.; Nugent, N.P.; Schweickhardt, R.L.; Rosenthal, J.L.; Overton
Mol. Cell. Endocrinol. 89, 141-151, 1992
A;Title: The cloning of the human follicle stimulating hormone receptor and its expression
A;Reference number: I57672; MUID:93246012; PMID:1301382
A;Accession: I57672
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-679, 'N', 681-695 <KEL>
A;Cross-references: GB:S59900; NID:G9300072; PIDN:AA26480.1; PID:G9300073
R;Minagishi, T.; Nakamura, K.; Takakura, Y.; Ibuki, Y.; Igarashi, M.
Biochem. Biophys. Res. Commun. 175, 1125-1130, 1991
A;Title: Cloning and sequencing of human FSH receptor cDNA.
A;Reference number: JN0122; MUID:91222171; PMID:1709010
A;Accession: JN0122
A;Molecule type: mRNA
A;Residues: 1-111, 'T', 113-136, 'AV', 199-306, 'A', 308-695 <MIN>
A;Cross-references: EMBL:M65085; NID:G182770; PIDN:AAA52477.1; PID:G182771
C;Genetics:
A;Gene: GDB:FSHR
A;Cross-references: GDB:127510; OMIM:136435
A;Map position: 2p21-2p16
A;Introns: 223/3
A;Note: the exact position of the intron cannot be determined from the experimental data
C;Function:
A;Description: receptor that mediates the biochemical effects of follitropin
A;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
F;1-695/Product: Follitropin receptor precursor, long splice form #status predicted <SPH
F;1-223,286-695/Product: follitropin receptor precursor, short splice form #status predi
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-695/Product: follitropin receptor #status predicted <MAT>
F;16-366/Domain: extracellular hormone binding #status predicted <EHB>
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;367-387/Domain: transmembrane #status predicted <TM1>
F;398-421/Domain: transmembrane #status predicted <TM2>
F;444-465/Domain: transmembrane #status predicted <TM3>
F;486-508/Domain: transmembrane #status predicted <TM4>
F;529-550/Domain: transmembrane #status predicted <TM5>
F;574-597/Domain: transmembrane #status predicted <TM6>
F;609-630/Domain: transmembrane #status predicted <TM7>
F;191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 13.2%; Score 621.5; DB 1; Length 695;
Best Local Similarity 22.4%; Pred. No. 1.7e-32;
Matches 207; Conservative 146; Mismatches 36; Indels 265; Gaps 25;
8 VLLSLPVLQLATGSSPRSVLLRGCPHCH---CEPDGRMLLRVDCSLGLSELPNL 64

Db 4 LLVSLAFLSLGSG-----CHHRI CHCSNRVFL---CQBSKVTEIPS-- 42
Qy 65 SVFTSYDLDNMNNSQLPPLPSLRFLELRAGNALTYPKGAFTGLVSL-KVLMLO 123
Db 43 -----DLPRNAT-----ELRFVLTKLRLVIQKGAFGSGFGLERIEISQN 80
Qy 124 NQLRHVTEALQNLRSLSLRLD-ANHISVVPSPCSGLSHSLRHLWLDNLTALTEIPQAF 182
Db 81 DVLVEIADVFSNLPKLUHEIRIEKANNLLYNP-----EAF 116
Qy 183 RSLSALQAMTLALNKHIIHPDYAFAGNLSLVVLHLHNNRIHSLGKCKGCDGLHSLTDLN 242
Db 117 QNLPLQYLLISNTGIKHLPDV-----HKHLSLQK-----VLLDIQ 152
Qy 243 YNNLDEPPTAIRLSNLKELGFHSNNRTSRIDEKAFVGNPSLTIITFDVNPQFVGRSAFQ 302
Db 153 DN-----INHTIERNISFVGLS-----FESVILWLNKNGIQ 183
Qy 303 HLPRLRLTLNGASQITEFPDLTGTANLESITLTGAQISSLPQVNCNLPNLQVLDLSYN 362
Db 184 ---EIHNCAPNG-----TOLDLNLUSD--NN 204
Qy 363 LLEDLPFSVCQKLIKIDLRHNEIYEIKVDTFQQLLSLSLNLAWNKIAIHPNAFSTLP 422
Db 205 NLEELPN-----DVFHGAS 218
Qy 423 SLIKLDLSSNLLSPFTTGLHGLTHLKLGTNHAQLSLSSENPFELKVIEMPYAYOCAP 482
Db 219 GPVILDISRTIRIHSLSYGLNKLKLRARSYTNLKKLPLEKLVAMLEASITSPSHCAF 278
Qy 483 G-----VCENAYKISNQ-----WNKGDNSMDDLHKDKAGFQAQDRDLDF 525
Db 279 ANWRQISELHPICNKS--ILROEVDVMTQGRQSSLAEDN-----ESSYRGFDMT 329
Qy 526 LLOPEEDL-KALHSVQCSPPGPKCEHLIDGLWLIIRIGVMTIATVLTALTCNALVTSTVFR 584
Db 330 YTEFDYDCNEVDVTCSPKPDAPNCPEDINGYNILAVLWIFISILAITGNIIVLVILT 389
Qy 585 SPLYISPKLLIGVIAVNMMLTGVSVAVLGADVAFPGSFARHGAWNGVGVCHVIGFLS 644
Db 390 SQYKLTVPRLMCLNAPADLCIGIYLLLIASVDIHTKSQYHNYAIDMQTAGCDAAGFE 449
Qy 645 IFSESSVFLTLAALRERGSVKYSAKFEKAPSSLSKVIILLCALLALTAAPVPLGGS 704
Db 450 VFASLSVYVLTALTLEKRWHTIHAMQDCKVQLRHASVVMWGWIFAFAPALPFIIGIS 509
Qy 705 KYGASPLCLPLPFGPESTMGMVALILLNSLCFLMMWTIATKLYCNLDKGLD-ENIWDCS 763
Db 510 SYMKVSICLPMDIDSPLSQLYVMSLVNLVLAFFVICGCIHIYLTVRNPNIVSSSDTR 569
Qy 764 MVKHALLLFTNCILNCPVAPLSPSSLINITFTSPEVVKILLVVVLPDACLNELLILF 823
Db 570 IAKRMAMLIPTDFLCMAPISPFAISAKLVPLITVSXAKILLVLFHPINSCANPLFAIF 629
Qy 824 NPHEKEDLVSLRKQTYVWTSKHPSLMSINSDDDVEKOSCDSTQALV--TFTSSSITVDLP 881
Db 630 TKNFRRRFFIL-----LSKCGCYEQMAQIYRTSTSTVHNTHP 667
Qy 882 PSSVSPSPAYPVTSCHLSSVAVFP 905
Db 668 RNHCSSAPRVTSQ---STYILVP 688

RESULT 11
A42395
lutropin receptor - mouse
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42395
R;Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.
J. Biol. Chem. 267, 4479-4488, 1992

A;Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenylyl

A;Reference number: A42395; MUID:92165799; PMID:1311310

A:Accession: A42395

A: Status: preliminary

A: Molecule type: mRNA

A:Residue type: **UNK**
A:Residue: **1-700 <GTD>**

A; References: 1-700 3002
A; Cross-references: UNIPROT:P30730; GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39432.1;
A; Note: sequence extracted from NCBI backbone (NCBIN:84064, NCBIP:84066)

C; Keywords: G protein-coupled receptor; transmembrane protein

F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F/153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F/180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match	12.8%;	Score 600;	DB 2;	Length 700;
Best Local Similarity	23.0%;	Pred. No. 4.3e-31;		

Matches 207; Conservative 143; Mismatches 320; Indels 230; Gaps 23; Gaps 23;

06
4 DUDAT DACT VIT AMT VIT KASOT HSDFT SCSSCDBEDFCACADG--ALDCCBDDPAGI AD- - 58
:
5 RGVVLSLPVLTLQATGSSSPSGVLTG-CPTHCHCEPDGRMLLRVDCSDLGLSELPSN 63
07

DD 4 RVPALLRQULVLAHLVLRQSLHSPFLSSGSRCPPEUCAPDG--ALUKCPGPRAGTAKL----- 58

QV 64 I.SVPTTSYINL.SMNNT SOLILPNBIDPSLREI.RELRIAGNAL.TVTPDKGAPETGIVS--IKUIMIQ 122

D _B	59	-----SUTYI-----PKKVTIPSOAFRGINEVVKIRISO 86
VY	64	LSVFISLSDSSMNATLSQULNFNFPFSUKFDEBLRAGNALITTFPAGAGTIGSIS-LKVLTJLQ 122

55 -----SHLL-----FVKVFQQARAGNENVAATISQ 88
 123 NNOLRHVPTALONLRISQSL-DANHSKVPSCSEGLSLRHLWLDDNAITEIPVA 181

101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612

QV 182 FRSLSALQAMTIALNKTHHIDPYAFGNLSLWTLHLNRRIHSLGKKCFDGLHSLETLDL 241

	-----CNCGRTLP--DV 145	: :
Db	135 -----	: :

Qy 242 NYNNLDEPTAIPRTLNLKEIGFHSNNIRSIPEKAFVG-NPSLITTHFYDNP IQFVGRSA 300

Db 146 SKISSENFILFICDNL-----YITTPGNAFQGNNESITLKLNGFEEVQSHA 197

QY 301 FQHLPELRLLTINGASOITEFPDLTGTANLESLLTGAQISSLPQTVTCNQLPNLQVLDLS 360

Db 198 F-----NGTTLIS-----LELK 209

Qy 361 YNL-LEDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLLSRLNLAWNKIAIHHNAPS 419

Db 210 ENIYLERKHS-----GTFQG-----A 225

Qy 420 TLP SLIKL DSS NLL SFP FTCLHG ITHL KLTGNHALQSL SSNF FFLKVIEMPYAYQC 479

Db 226 TGPsi--LDVSSTKLQALPSHGLESIQTLTIATSYSLKTLPSREKFTSLIATLTPSHC 283

QY 480 CAFGVCN-----AYKTSNQWKNKGDNSSMDLHKKD--AGMFOAQDERDLEDFLLDFFED 532

Db 284 CAFRNLPKKEQNF SFSIFENFSKQCESTVREANNETLSAIF---EENELSGW---DYDYD 338

Qy 533 LKALHSVQCSPPGPFKCEHLLDGWLIRIGVWTIAVLATCNALYTSTVPSRPLYISPI 592

Db 339 FCSPKTLQCTPEPDFAFNPCEDIMGYAPLRVLINILIAIFGNLTWLFVLLTSRYKLTVP 398

Qy 593 KLLIGVIAVMMLTGVSSAVLAGVDAFTFGSPARHGAWENGVGCHVIGTISIPASESV 652

Db 399 RFLMNLSPADFQWGLLYLLIASVDSQTKQYYNHAIQWGTSGCSAAGFTTVFASLSV 458

653 FLITLALERGPSVKYSAKFETKAPSSLVKIIICALLALTWAAVPLLGSKYGASPLC 712

DB 459 YTLAVTLERWHRTTAYQLDQDKRHRHAPIMLEGMIFSTLMATPLVGVSSYKVKV518
713 13 DECEMBER/04/08 11 11 NOT GET MUMTAVMST VOMT DV/04/08 ENTIDOC000000011111 721

QY 713 LPLPFGEBSWGMYVAILLNSLCFLMNTIATKLYCNDRGDL-ENIWDCSMVRIALL 771

Db 88 -----TLORLESFYNLSKMTHEIRNTR-----SLTSD----- 118
QY 248 EPTAIRTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSAFOHLPEL 307
Db 119 --PDALKELPLKPLGIGTNGVFPDVKV-----YSTDVFI----- 155
QY 308 RLTTLNGASQITEPDLTGANTLESLLTGAQISLP-----QTVCNQLPNQLVDLSYVL 363
Db 156 -----LEITDNP-----YMASIPANAFQGLCNELTTLKL-----YN- 186
QY 364 LEDLPSFSVCOKLIDLRHNEIYEIKVDTFQOLLSLRSLNLANWK-TAIIHPNAPSTL- 421
Db 187 -----NGFTSIQGHAFNG-TKLDAYVNLKKNKYSATDKDAFGVY 225
QY 422 --PSLIKLDLSNLSASPITIGLHGLTHLKTGNHALQSLISSENPPKLVIEPMPYAYQC 479
Db 226 SGPTL--LDVSYTSVTALPSKLEHLKELIARNTWTAKKPLSLSLFLHLTRADLSYPSHC 283
QY 480 CAP-----GVCE-----NAYKISNOWNKGDNSMDLHK--KDAQMP 514
Db 284 CAFKNQKIRGILESMLCNESIRSIRQRKSVNTLNGPFDQEBEYELGDSHAGYKNSQF 343
QY 515 QAQDE-----RLEDFLDF-----BEDLKALHS-----VOCSPSPG 546
Db 344 QDTSNSHYVFFERQDEILGFGQELKNPOBETQAFDSDHYDVTVCNGNEDMVCPTKSD 403
QY 547 PFKPCEHLLDGWLIRIGVTTAVALTCTNALVTSTVFRSPLYISPIKLLIGVIAAVNMLT 606
Db 404 EFNPCEDIMGYKFLRIVVFWVSLALLGNVFLVLLTSHYKLTVPFLMCLNAPADFCLM 463
QY 607 GVSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIPASESSVFLTLTAAALGRFSV 666
Db 464 GMYLLIASVDLYTHSEYNNHAIDQGTGPGCNTAGFTVFASLSVYTLTITLERWYAI 523
QY 667 KYSAFETKAPFSSLVII-----LLCALLATMAAVPILGSKYKASPLCLPLPGEPEST 722
Db 524 TFAMLRDRIRURHAYAVMGWVCCFLAL-----LPLVGISSYKAVKVICLPMDETETPLA 579
QY 723 MGYVALLNLSCLFMMTIATYKLYCNL-----DKGDLNIWDCSMVKHIALLLFTNCI 777
Db 580 LAYIILVLLNVAIIVCSCVKIYITVRNPQYNGDK-----DTKIARMAVLIFTDFM 635
QY 778 LNCVAFISFSLINLTIPISEVIFKILLVVVPLPACLNPLLYILFNPHFKEDLVSL--- 834
Db 636 CMAPISFYALSALMKPLITVNSKILLVFLVPLNSCANPFLYAFTRAFORDVPIILSK 695
QY 835 ----RKQTVVTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSITDYDLPSSVPSPAY 890
Db 696 FGICKRQAOAY-RGQVSPKNSAGIQIKVTRDMRQSLPNMODE-----Y 739
QY 891 PVTESCHLS 899
Db 740 ELLENSHLT 748

RESULT 14

luteinizing hormone/chorionic gonadotropin receptor - rat
C:Species: Rattus sp. (rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: I77463
R:Aatsinki, J.T.; Pietila, E.M.; Lakkakorpi, J.T.; Rajaniemi, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A:Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an
A:Reference number: 157668; PMID:92347604; PMID:1353463
A:Accession: I77463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-700 <RES>
A:Cross-references: GB:S40803; NID:g252163; PIDN:AAB22680.1; PID:g252164
C:Genetics:
A:Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat H

F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 12.5%; Score 588.5; DB 2; Length 700;
Best Local Similarity 22.6%; Pred. NO. 2.4e-30;
Matches 205; Conservative 143; Mismatches 309; Indels 251; Gaps 25;

QY 6 LGVLLSPVLVQLATGSSPSRSGVLLRCPCPHCHCEPDGRMLLRVDCDGLGSELPSNLS 65
Db 14 LAVLLKPSQLQ-----SRELSG---SRCPEPCAPDG--ALRCFGRAGLARL----- 58
QY 66 VFTSYLDLSMNNISQLLPNPLSRFLFEELRAGNALTYIPKGAFTGLYS-LKVLMLQNN 124
Db 59 -----SITYL-----PVKVIPOAPRGLNEVVKIISQSD 88
QY 125 QLRHVPTAQLNQLSLOSLRL-DANHSYVPPSCFSGHLSRHLWLDNALTEIPVQAFR 183
Db 89 SLERIANAPDNLANLSSELLIONTKNLLYIEFGAFTNLPRLKYLSI----- 134
QY 184 SLSALQMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLETLDNY 243
Db 135 -----CNTGIETLP--DVTK 147
QY 244 NLDSEPTAIRTLSNLKELGPHSNIRSIPEKAFVG-NPSLTIHFYDNP IQVGRSAFO 302
Db 148 ISSSEFNILSICNL-----HITIPGNAFOGMNESVTCLKYNGEVEVQSHAF- 198
QY 303 HLPRLTTLNGASQITEPDLTGANTLESLLTGAQISLPQTVCNQLPNQLVDLSYN 362
Db 199 -----NGTTLIS----- 205
QY 363 LLEDLPFSVCOKLQKIDLRHNEIYEIKVDTFQOLLSLRSLNLANWKTAIIHPNAP--ST 420
Db 206 -----LELKEN-IYLEK-----MHSGAFOGAT 226
QY 421 LPSLIKLDLSNLSRPPITGLHGLTHLKTGNHALQSLISSENPPKLVIEPMPYAYOC 480
Db 227 GPSI--LDISSTKQALPSHGLSIOQLIALSSVSLKTLPSKEKFTSLVATLTPSHCC 284
QY 481 A-----FGVCENAYK-ISBNOWNKGDNSMDLHKDKAGMFOAORDERLDFLL 527
Db 285 AFRNLPKKEQNFSPSIPENFSKQCESTVRKADNETL-----YSAIF--EENELSGW-- 333
QY 528 DFEEDLKALHSVQCSPPSPGPKCEHLDDGWLIRIGVTTAVLTCTNALVTSTVFRSPL 587
Db 334 DYDYGFCSPKTLQCAPEDAFNPCEDINGYAFRLVLIWLINILAFGLNLTVLVLLTSRY 393
QY 588 YISPIKLLIGVIAAVNMLTGVSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFA 647
Db 394 KLTVPFLMCLNSFADFCMGYLLILLIASVDSQTKQYTNHAIDMTQSGCGGAGFTVFA 453
QY 648 SESSVFLITLAAALBERGFSVKYSAPFETKAPFSSLSKVIILLCALLATMAAVPLLGSKY 707
Db 454 SELSVYTLITVITLERWHITITVAVQDQKLRLHAIPMLGGWLFTSLIATMPLVGSINVM 513
QY 708 ASPLCLPLPGEPESTMGVNWALILNLSCLPLMTIATYTKLYCNLDKGDL-ENIWDSCSVK 766
Db 514 KVSICLPMDEVSTLSQVYLSILINLVAVFVVCACYIRIYFAVQNPBELTAPNKDTKAK 573
QY 767 HIALLTNCLNCPVAPLSPSSLINLTFSPEVIFKILLVVVPLPACLNPLLYILFNPH 826
Db 574 KMAILIFTFTCMAPISFAISAPKPVLTITVNSKILLVFLVYPVNSCANPFLYAIPTKA 633
QY 827 FKEDLVSL-----RKQTVVTRSKHPSLMSINSDDEKQSCDSTQALVTFTSSSIT 877
Db 634 FORDFLLLSRFGCKKRAELY---RRKEFSAYTSNCKNGPPGASKPKSQATLKLSTVHCQ 690
QY 878 YDLPPSSV 885
Db 691 QPIPPRAL 698

RESULT 15
A49744

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:53:17 ; Search time 9805 Seconds
(without alignments)

4482.293 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTESCHLSVAFVPCPL 907

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cpn2_1/USPTO.spool/US10751736/runat_12072005_083640_22171/app_query.fasta_1.1095
-DB=GenEmbl -QFWT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANSHUMAN40.cdi -LIST=45
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -WAXLEN=2000000000
-DICALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-USER=US10751736 @CGN 1 1.6632 @runat_12072005_083640_22171 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.roi.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	2724	6	CQ721366 Sequence
2	907	100.0	2724	6	CQ766844 Sequence
3	907	100.0	2724	6	AR308745 Sequence
4	907	100.0	2724	6	AX549136 Sequence

5	907	100.0	2880	6	AX658241 Sequence
6	907	100.0	2880	9	AF062006 Homo sapi
7	907	100.0	4559	6	BD075815 G protein
8	907	100.0	4570	6	AR447625 Sequence
9	820	90.4	2724	6	AR308754 Sequence
10	730	80.5	2724	9	AF061444 Homo sapi
11	606	66.8	2082	6	BD135244 Novel mam
12	362	39.9	145165	9	AC078860 Homo sapi
13	357	39.4	4095	9	AK075399 Homo sapi
14	202	22.3	606	6	E42915 Novel gonad
15	202	22.3	606	6	AX016185 Sequence
16	71	7.8	162004	9	AC090116 Homo sapi
17	51	5.6	3115	10	AF110818 Mus muscu
18	44	4.9	816	11	BV208561 GPR49_209
19	32	3.5	176092	10	AC129329 Mus muscu
20	32	3.5	195389	2	AC126943 Mus muscu
21	32	3.5	222154	2	AC126127 Rattus no
22	32	3.5	239603	2	AC132780 Rattus no
23	29	3.2	195389	2	AC126943 Mus muscu
24	29	3.2	252477	2	AC136417 Rattus no
25	27	3.0	443	4	AF455789 Sus scro
26	22	2.4	350	5	CR352678 Gallus ga
27	19	2.1	60	6	CQ550152 Sequence
28	18	2.0	284	6	E42919 Novel gonad
29	18	2.0	284	6	AX016194 Sequence
30	18	2.0	1827	6	CQ842816 Sequence
31	18	2.0	1827	9	AK123055 Homo sapi
32	18	2.0	2208	6	AX301830 Sequence
33	18	2.0	2484	9	AF190501 Homo sapi
34	18	2.0	2487	6	AX926297 Sequence
35	18	2.0	2711	6	AX301828 Sequence
36	18	2.0	2901	6	AX301824 Sequence
37	18	2.0	2901	6	AX301833 Sequence
38	18	2.0	3041	6	E42917 Novel gonad
39	18	2.0	3041	6	AX016191 Sequence
40	18	2.0	3119	6	AX549295 Sequence
41	18	2.0	3119	9	AB049405 Homo sapi
42	18	2.0	3273	6	CQ870634 Sequence
43	18	2.0	3286	9	BC047905 Homo sapi
44	18	2.0	3306	6	CQ840811 Sequence
45	18	2.0	3325	9	AY358119 Homo sapi

ALIGNMENTS

RESULT 1
CQ721366
LOCUS CQ721366 2724 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 7300 from Patent WO02068579.
ACCESSION CQ721366
VERSION CQ721366.1 GI:42282223
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 7300 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source
1..2724
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2724
Score: 907.00 Matches: 907

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-751-736-84 (1-907) x CQ721366 (1-2724)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB 1 ATGGAGACCTCCCGGCTCGGTGTCTCTCTTGGCTGTGCTGCTGAGCTGGCGACC 60

QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB 61 GGGGCGAGCTCTCCAGGCTCTGGTGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 120

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB 121 GAGCCGAGCGGAGGATGTGCTCAGGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180

QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB 181 CCTTCCAACTCAGCGTCTTCACTTCCCTACCTAGACCTCAGTATGAACAAATCAGTCAG 240

QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
DB 241 CTGCTCCCGATCCCTCCCGAGTCTCCGCTTCTTGAGGAGTTACGTCTTGGCGGAAAC 300

QY 101 AlaLeuThrTyrIleProGlyGlyAlaPheThrGlyLeuTyrSerLeuValLeuMet 120
DB 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTTATG 360

QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB 361 CTGCAGAAATAATCAGCTAAGACACACTACCCACAGAAAGCTCTGCAGAAATTTGCAAGCCTT 420

QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB 421 CNAATCCCTGCGTCTGGATGCTTAACACATCAGTATGTGCCCCCAAGCTGTTTCAGTGGC 480

QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
DB 481 CTGCATTTCCCTGAGGCACCTGTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540

QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIleHisHis 200
DB 541 GCTTTTAGAAGTTATCGGCATTGCAAGCATGACCTTGGCGGCTTGAACAAATAACACCAC 600

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
DB 601 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660

QY 221 ArgIleHisSerLeuGlyIlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720

QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 721 TTAATAATCAATAAACCCTTGATGAATTTCCCACTGCAATTAGAGACTCTCCAACCTTAAA 780

QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 781 GNACTAGAAATTCATAGCAACAAATATCAGTCTGATACCTTGAGAAAGCAATTTGTAGGCAAC 840

QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 841 CCTTCTCTATTACAATAATCTTATGACAATCCATCCAATTTGTTGGGAGATCTGCT 900

QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 901 TTTCAACATTTTACCCTGAACACTAGAACACTGACTCTGAATGGTGGCTCAAAATTAACGAA 960

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340

DB 961 TTTCTCTGATTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC 1020

QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTACTTAATCTCCAAGTCTAGATCTGTCT 1080

QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAAGATTACCCAGTTTTTTCAGTCTGCCAANAAGCTTCAGANAATTGAC 1140

QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
DB 1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTGTCTTAGCCTC 1200

QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTCCTATTATTACCCCAATGSCATTTTCCACT 1260

QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGCCATCCCCTAATAAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTATAAAT 1320

QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGCTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380

QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGCTGTGT 1440

QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGAGTGTGTGAGATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500

QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGAGGACCTTTCATAAGAAAGATCTCGAATGTTTCAAGGCTCAAGATGAACGT 1560

QY 521 AspLeuGluAspPheLeuLeuAspPheGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGGCCCTTCATTCAGTGGAG 1620

QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCGAGGCCCTTCAACCTCTGGAACACCTGCTGTGATGGCTGCTGATC 1680

QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATTC 1740

QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACACTGTTAAATGGGGTTCATCGCA 1800

QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGGCTGGTGGATGGCTTCACT 1860

QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGGAGCTTTGGACGACATGGTGGCTGGTGGGAGAAATGGGTGGTGGTGGTGGTGGTGGT 1920

QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
DB 1921 GGTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTCTGCTTACTCTCGCAGCCCTG 1980

QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGTGGTGTCTCTGTGAATAATTTTCGAAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040

QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2041 CTGAAAGTAAATCAATTTTGTCTGTGGCTGTGGCTGTGACCATGGCGCGCAGTTCCTCCCTG 2100

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTTCTCAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTCAGTGTGCCAAAGCTTCAGAAAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAGACATAATGAATCTACGAATTTAAAGTTGACACTTTCCAGAGTGTGTTAGCCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTGCTATTATTCACCCCAATGCAATTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGGCATCCCTAATAAGCTGACCTATCGTCCAACCTCTGCTCGCTTTTCCCTATAACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGAAATCATGGCTTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGGAGTGTGAGAAATGCCATATAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGAGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATAGAGTCTGGCACCTTACTTGTAACTTTGTGGTACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAATCTGTAATGGGGTCAATGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGCGTGTGGCTGGTGTGGATCGTTCACT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTGGCAGCTTTGGCAGCATGGTGTGCTGGGAGAAATGGGGTGGTTGGCATGTCAATT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTGTCTCCATTTTCTTCCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700

DB 2041 CTGAAGTAATCAATTTTGTCTGTGTGCTGTGTGGCTTGACCATGGCGCAGCTTCCCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLysLeuProLeuProPheGlyGluPro 720
DB 2101 CTGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGTGCTGCCTTTGCCCTTTTGGGGAGCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
DB 2161 AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
DB 2221 ACCATTGCTACACCAAGCTCTACTGCATTTGGACAGSGAGAGCTTGAGAAATATTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2281 GACTGCTCTATGGTAAAAACACATTGGCTGTGTCTTCCACCACTGCATCTCAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2341 CCGTGGGCTTTCTTGTCTCTTCTCTTTAATAAAACCTTACATTTATCAGTCTCGAAGTA 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2401 ATTAAGTTTATCTCTCTGCTGGTAGTCCACTTCTGCTGATGCTCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2461 ATCTGTGTTCAATCTCTCACTTTAAAGGAGGATCTGGTGAGCTCGAGAAAGCAAACTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
DB 2521 TGGACAGATCAAAACACCCCAAGCTTGATGTCATTAATCTCTGATGATGTCGAAAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2581 TCCTGTGACTCACTCAAGCCTTGTAACCTTTACCAGCTCCAGCATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2641 CTTCCAGTTCGTTGCCATCACAGCTTATCCAGTGAAGTGAAGTGCATCTTTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCAATTTGCCCATGTCTC 2721

RESULT 3
LOCUS AR308745 2724 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 263 from patent US 6555339.
ACCESSION AR308745
VERSION AR308745.1 GI:31700274
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2724)
AUTHORS Liaw,C.W., Behan,D.P. and Chalmers,D.T.
TITLE Non-endogenous, constitutively activated human protein-coupled receptors
JOURNAL Patent: US 6555339-A 263 29-APR-2003;
FEATURES Location/Qualifiers
1..2724
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2724
Score: 507.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0


```
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTTCCCTCAAAACCGCTGCAATCAGTTACCTAAATCTCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAGATTATCCAGTTTTTCNGTCTGCAAAAGCTTCAGAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCAGCAGTTGCTTAGCCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
QY 421 LeuProSerLeuLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
DB 1261 TTGCCATCCCTAATAAAGCTGGACCTATGTCACACCTCTCTGTGCTTTTCTATACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTCTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGGAGTGTGAGAAATGCTATAGAATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1560
QY 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTCTCGTCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCAGGCCCTTCAAAACCTGTGAACACTGCTGTATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATTGGAGTGTGACCATAGCAGTCTTGCCACTTACTTGAATGCTTTGGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGGTTCATGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 CAGATGAACATGCTACGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGGTTCCTACT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGGCAGCTTTGGCAGCATGGTGGCTGGGAGATGGGGTGGTGGCTGATGCTATT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2041 CTGAAAGTAAATCATTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
```

```
DB 2101 CTGGGTGGCAGCAGATATGGCCCTTCCCTCTCTGCTGCTTTCCTTTTGGGGAGCCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
DB 2161 AGCACCATTGGGTACATGGTCTCTCATCTTCTCAATTCCTTTCCTTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
DB 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTTGGACAGGGAGACCTCGAGAAATATTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2281 GACTGCTCTATGGTAAACACATTCCTGTTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2341 CTTGTGGCTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2401 ATTAAGTTTATCTCTCTGCTGCTAGTCCCACTCTCTGCTGCTCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2461 ATCTTGTTCATTCCTCTTAAAGGAGATCTGGTGAGCTCGAGAAAGCAAACTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
DB 2521 TGGACAGATCAAAACACCCAGCTTGATGTCATTAATCTCTGATGATGTCGNAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2641 CCTCCAGTTCGTCCTCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCAATTGTCCCATGTCTC 2721

RESULT 5
LOCUS AX658241 2880 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 157 from Patent WO03000928.
ACCESSION AX658241
VERSION AX658241.1 GI:29160776
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
Petersen,M.W. and Elsner,H.I.
TITLE Methods for identification of cancer cell surface molecules and
cancer specific promoters, and therapeutic uses thereof
JOURNAL Patent: WO 03000928-A 157 03-JAN-2003;
Odin Medical A/S (DK)
FEATURES
Location/Qualifiers
source 1..2880
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```


Db 589 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGCCCTGAAACAAAATACACCAC 648
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn 220
Db 649 ATACCAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCCATACCAAT 708
Qy 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 709 AGAATCCACTCCCTCGGGAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 768
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 769 TTAATAATACAAATAACTGATGAATTCGCCACTGCAATTTAGGACACTCTCCAACCTTAAA 828
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 829 GAACTAGGATTTTCATAGCAACAATATCAGGTGCTGATACCTGGAAGAAGCAATTTGTAGGCAAC 888
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 889 CTTCTCTTATTAACAATACATTTCTATGACAATCCATCCCAATTTGTTGGGAGATCTGCT 948
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 949 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTTCACAAATAACTGAA 1008
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1009 TTTCTGATTTTAACCTGGAACCTCAACCTGGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1068
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1069 TCATCTCTTCCTCAAAACCTCTGCAATCAGTTACCTAATCTCCAAGTCTAGATCTGCT 1128
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1129 TACAACCTATTAGAAGATTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1188
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1189 CTAAGACATAATGAATCTACGAAATTAAGATTGACACTTTCCAGCAGTTGCTTAGCCCTC 1248
Qy 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db 1249 CCAATGCTGCAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTTCCACT 1308
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuLeuSerPheProIleThr 440
Db 1309 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACCTCCTGCTGCTTTTCTATAACT 1368
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1369 GGGTTTACATGGTTAACTCACTTAATAATTAACAGGAATCATGSCCTTACAGAGCTTGATA 1428
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1429 TCATCTGAATAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTAGCTTACCAGTGTGT 1488
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1489 GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1548
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1549 AGCAGTATGCGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1609 GACCTTGAGATTTCTGTTGACTTTGAGGAAGACCTGAAGACCTTCATTCAGTGCAG 1668
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1669 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTAACACCTGCTTGAACACCTGCTTGAATGGCTGGCTGATC 1728

Qy 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1729 AGAATTTGGAGTGTGACCATAGCACTTCTGCACTTACTTGTATATGCTTTGGTGCATTCA 1788
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1789 ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCAATTAACCTGTTAATTTGGGTCAATGCA 1848
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1849 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCGCTGGCTGGTGTGGATGCGTTCACT 1908
Qy 621 PheGlySerPheAlaAaTqHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTGGCAGCTTTGACGACATGGTCCCTGGTGGGAGAAATGGGGTTGGTTCGCAATGTCATT 1968
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1969 GGTTTTTTGTCCATTTTGTTCAGAAATCATCTGTTTTCTTCTTACTCTGGCAGCCCTG 2028
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2088
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAGTAATCATTTTGTCTGTGCCCTGTGGCTTGACCATGGCCGAGTTCCTCTG 2148
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2208
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuAsnSerLeuCysPheLeuMetMet 740
Db 2209 AGCACCATGGGCTACATGGTGCCTCTCATCTTGTGCTCAATCCCTTGTCTCTCATGATG 2268
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATTGGCTACACCAAGCTCTACTGCATTTTGGCAAGGAGACCTGGAGAAATATTGG 2328
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2329 GACTGCTCTATGGTAAAAACACATTTGCCCTGTGCTCTTCCACCACTGCATCTCAAACTGC 2388
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2389 CCTGTGGCTTTTCTTGTCTCTCTCTTTAATAAACCTTATCATTTTATCAGTCCCTGAAGTA 2448
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2449 ATTAAAGTTTATCTCTGTGGTAGTCCCACTTCTCGCATGCTCAATCCCTCTCTCTAC 2508
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2509 ATCTTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGAGCTTGAGAAAGCAAACTACGTC 2568
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2569 TGGCAAGATCAAAACACCCCAAGCTTGTATGCAATTAATCTGTATGATGTGAAAAACAG 2628
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2629 TCCTGTGACTCACTCAAGCCTTGGTAACCTTTACCAGCTCCAGCATCACTTATGACCTG 2688
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2689 CCTCCAGTTCGGTGCCATCACCACTTATCCAGTACTGAGAGCTGAGAGTGCATCTTCTCTCT 2748
Qy 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCAATTTGTCCCATGTCTC 2769

[illegible]

Db 1689 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGAAATTAAGGTGACAAAC 1748
Qy SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1749 AGCAGTATGGACGACCTTCATAAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1808
Qy AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1809 GACCTTGAGATTTCTCTGCTGACATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1868
Qy CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuile 560
Db 1869 TGTTTCACTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC 1928
Qy ArgileGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1929 AGAATTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTTGGTGAATTC 1988
Qy ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuileGlyValIleAla 600
Db 1989 ACAGTTTTTTCAGATCCCTCTGACATTTCCCCCATTTAAATGGGGTCAATCGCA 2048
Qy AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2049 GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGTGTGATGGCTTCACT 2108
Qy PheGlySerPheAlaArgHisGlyAlaTyrPheGluAsnGlyValGlyCysHisValIle 640
Db 2109 TTTGGCAGCTTTGGCAGCATGGTGGCTGGTGGGAGAAATGGGTTGGTTGCCATGTCATT 2168
Qy GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrIleuAlaLeu 660
Db 2169 GGTTTTTTGTCTCATTTTGTCTCAGAAATCATCTGTTTCTCTGCTTACTTGGCAGCCCTG 2228
Qy GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2229 GAGCGTGGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2288
Qy LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2289 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2348
Qy LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2349 CTGGTGGCAGCAAGTATGGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2408
Qy SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2409 AGCACCATGGGCTACATGGTCTCTCATCTTGTCTCAATTCCTTTTGTCTTCTCATGATG 2468
Qy ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyrp 760
Db 2469 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAATATTTGG 2528
Qy AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2529 GACTGCTTATGGTAAAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2588
Qy ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2589 CTTGTGGCTTCTCTGCT 2648
Qy IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2649 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708
Qy IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2709 ATCTTGTTCATTCCTCCTTAAAGAGGATCTGGTGAGGCTCTGAGAAAGCAACCTACGTC 2768
Qy TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860

Db 2769 TGGCAAGATCAAAACACCCCAAGCTTGATGTCATTAATTAATCTGATGATGTCGAAAAACAG 2828
Qy SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrThrAspLeu 880
Db 2829 TCCTGTGACTCACTCAAGCTTGGTAACTTACCTTTTACAGCTCCAGCATCACTTATGACCTG 2888
Qy ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2889 CTTCCAGTTCGGTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2948
Qy ValAlaPheValProCysLeu 907
Db 2949 GTGGCAITTTGTCCTCATGTC 2969
RESULT 8
AR447625 4570 bp DNA linear PAT 20-FEB-2004
LOCUS Sequence 201 from patent US 6673549.
DEFINITION
ACCESSION AR447625
VERSION AR447625.1 GI:42675949
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4570)
AUTHORS Furness,L.M. and Buchbinder,J.L.
TITLE Genes expressed in C3A liver cell cultures treated with steroids
JOURNAL Patent: US 6673549-A 201 06-JAN-2004;
FEATURES
Location/Qualifiers
1. 4570
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4570
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-751-736-84 (1-907) x AR447625 (1-4570)
Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuAlaThr 20
Db 262 ATGGACACTCCCGCTCGGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 321
Qy 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 322 GGGGGCAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACATGTCATGTC 381
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 382 GAGCCGACGCGCAGCATGTTGCTCAGGGTGACTGCTCCGACCTCGGGGCTCTCGGAGCTG 441
Qy 61 ProSerAsnLeuSerValPheThrSerTyrIleAspLeuSerMetAsnAsnIleSerGln 80
Db 442 CTTTCAACCTCAGCGTCTTCACTCTTACCTAGACCTCAGTAAGAAACAAATCATAGT 501
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
Db 502 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTGAGGAGGATACGCTTTCGGGGAAAC 561
Qy 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 562 GCTCTGACATACATTCACCAAGGAGCATTCATCGGCCCTTTTACGCTTTAAAGTTCTTATG 621
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 622 CTGCAGAAATAATACGTAAAGACACGTACCCACAGAGGCTCTGCAGAAATTTGGAGGCTT 681
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160

Db 682 CAATCCCTGGCTGGATGCTAACCACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 741
Qy 161 LeuHisSerLeuArgHisLeuTyrLeuAspAsnAlaLeuThrGluIleProValGln 180
Db 742 CTGCATTCCCTGAGGCACCTGTGGCTGGATGACAATGGGTTAACAGAAATCCCCGTCAG 801
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db 802 GCTTTTAGAAGTTTATCGGCATTTCGCAAGCCATGACCTTGGCCCTGGAACAAAATACACCAC 861
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 862 ATACCAGACTATGCCCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTCCCAATAACAAT 921
Qy 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 922 AGAATCCACTCCCTGGGAAGAAATGCTTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT 981
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 982 TTAATTTACAATTAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACCTTAAA 1041
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 1042 GAACCTAGGATTTTCATAGCAACAATATCAGTCCATACCTGAGAAAGCATTTGTAGGCAAC 1101
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1102 CCTTCTCTTATACATAACATTTCTATGACATCCCATCCATTTGTTGGAGACTGCT 1161
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1162 TTTCAACATTTACCTGAACTTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 1221
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1222 TTTCTCGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1281
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1282 TCATCTCTTCTCAACACCTGTCGATTCGATTCAGTTACTTAATCTCCAGTGTAGATCTGTCT 1341
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1342 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAAAATTGAC 1401
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1402 CTAAGACATAATGAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTTGCTTAGCCTC 1461
Qy 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1462 CGATCCCTGAAATTTGGCTTGGAAACAATAATGTCTATTATTCACCCCAATGCAATTTTCCACT 1521
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
Db 1522 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACCTCTCTGTGCTTTTCTCTATAACT 1581
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1582 GGGTTACATGGTTTAACTCATTAAATTTAAACAGGAAATCATGCTTATACAGAGCTTGATA 1641
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1642 TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCCCTTATCCACAGTCTGT 1701
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnIleProAsnLysGlyAspAsn 500
Db 1702 GCATTTGGAGTGTGAGAAATGCTATAGATTTCTTAATCAATGGAATAAAGGTGACAC 1761
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520

Db 1762 AGCAGTATGAGCAGCCTTCATTAAGAAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1821
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1822 GACCTTTGAAGATTTCCTGCTTTCAGCTTTGAGGAAGACCTCGAAAGCCCTTCATTCAAGTCAG 1881
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Db 1882 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATC 1941
Qy 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1942 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAACTGTTTGGTGAATTCA 2001
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 2002 ACAGTTTTTCAGATCCCTCTGTACATTTCCCACTTAAACTGTTAATGGGGTCATCGCA 2061
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2062 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGCTGGTGTGATGCGTTCACT 2121
Qy 621 PheGlySerPheAlaArgHisGlyAlaTyrTyrGluAsnGlyValGlyCysHisValIle 640
Db 2122 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCATT 2181
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 2182 GGTTTTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 2241
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2242 GAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTTGAACCGAAAGCTCCATTTTCTAGC 2301
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2302 CTGAAGATTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2361
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2362 CTGGGTGGCAGCAGATATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2421
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2422 AGCACCATGGGTACATGCTGCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2481
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
Db 2482 ACCATTGCCCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG 2541
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2542 GACTGCTCTATGGTAAACACATTTGCCCTGTTGCTCTTCCCACTGCATCTCTAAACTGC 2601
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2602 CCGTGGCTTCTTGTGCTTCTCCTCTTAAATAAACCTTACATTTATTCAGTCTGCTGAAGTA 2661
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2662 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2721
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2722 ATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGGTGGGCTTGAGAAAGCAACCTACGTC 2781
Qy 841 TrrThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2782 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTCTGATGATGCGAAAAACAG 2841
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2842 TCCTGTGATCAACTCAAGCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2901

QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2902 CCTCCAGTTCGGTCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCT 2961
QY 901 ValAlaPheValProCysLeu 907
Db 2962 GTGGCAATTGTCCCATGCTC 2982

RESULT 9

AR308754 LOCUS 2724 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 277 from patent US 6555339.
ACCESSION AR308754
VERSION AR308754.1 GI:31700283
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1. (bases 1 to 2724)
Liaw,C.W., Behan,D.P. and Chalmers,D.T.
Non-endorogenous, constitutively activated human protein-coupled
receptors

JOURNAL

Patent: US 6555339-A 277 29-APR-2003;

FEATURES

Location/Qualifiers
1..2724
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2724
Score: 820.00 Matches: 906
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 90.41% Indels: 2
DB: 6 Gaps: 0

US-10-751-736-84 (1-907) x AR308754 (1-2724)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 1 ATGGACACCTCCCGGCTCGGTGCTCCTGCTCTTGGCTGCTGCTGCTGCGGACC 60
QY 21 GlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 61 GGGGGCAGCTCTCCAGGCTCGGTGCTGCTGAGGGGCTGCCACACACTGTCTTGC 120
QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 121 GAGCCCGACGGAGGATGTGCTCAGGGGGGACTGCTCCGACTCGGGGCTCTCGGAGCTG 180
QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 181 CTTTCAACCTCAGGCTCTCACTTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
Db 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTTACGTCTTGGGGAAAC 300
QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTCATCTGGCCCTTTACAGTCTTAAAGTTCTTATG 360
QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 361 CTGCAGAAATATCAGCTATAGACACGTACCCACAGAGCTCTGCAGAAATTTGGGAGCCTT 420
QY 141 GlnSerLeuArgLeuAspAlaLeuHisIleSerTyrValProProSerCysPheSerGly 160
Db 421 CAATCCCTCGCTGAGTGTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180

Db 481 CTGATTCCTCGAGGACCTCTGGCTGGATGACAATGCTTTAAACAGAAATCCCGCTCCAG 540
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Db 541 GCTTTTAGAGCTTATCGGCATTGCAAGCCATGACCTTGGGCCCTTGAAACAAATATACAC 600
QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCNTAACAT 660
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 661 AGAATCCACTCCTCGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 721 TTAATTTACAATAACCTTTGATGAATTCCTCAATTCGCAATTAGGACACTCTCCAACCTTAA 780
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 781 GAACCTAGGATTTCTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 841 CTTTCTCTTATTAACAATACATTTCTATGCAATCCATCCCAATTTGTTGGGAGATCTGCT 900
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCTCAATTAACCTGAA 960
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyValGlnIle 340
Db 961 TTTCTCTGATTTAACTGGAACTGCAACCTGGAGAGCTGACTTTTAACTTGGAGCACAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAAACGCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1081 TACAACCTATTAGAGATTTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAGACATAATGAATCTACGAAATTAAGTTGACCTTTCCAGCAGTGTGCTTAGCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGACAAAATTCCTATTATTACCCCAATGCAATTTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAAATAAGCTGGACCTATCGTCCCAACCTCTGCTGCTCTTTTCTATAACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCAGTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAAACCTTCCAGAACTCAAGGTATAGAAATGCCTTATGCTTTACCAGTCTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAAATCCCTTATAAGATTTCTTAATCAATGGAATAAAGGTGACAC 1500
QY 501 SerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGGACGACTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGACGT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540

Db 1561 GACCTTGAAGATTTCCTGCTGACTTTGAGGAGACCTGAAGCCCTTCATTCACTGTCAG 1620
 Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuile 560
 Db 1621 TGTTCACCTTCCCGAGCCCTTCAAAACCTGTGAACACCTGCTTGCATGGCTGATC 1680
 Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
 Db 1681 AGAATTGGAGTGGACCATAGCAGTTCTGGCACCTTACTTGTAACTTTGGTGACTTCA 1740
 Qy 581 ThrValPheArgSerProLeuTrpIleSerProIleLysLeuLeuIleGlyValIleAla 600
 Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCATTAACCTGTAACTGGGGTCATCGCA 1800
 Qy 601 AlaValAsnMetIleuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
 Db 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGCGTCTGGCTGGTGGTGGTGGTCACT 1860
 Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
 Db 1861 TTTGGCAGCTTTTGCACGACATGCTGCTGGTGGGAGAATGGGGTTGGTTGCCATGTCAT 1920
 Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
 Db 1921 GGTITTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
 Qy 661 GluArgGlyPheSerValLysTrpSerAlaLysPheGluThrLysAlaProPheSerSer 680
 Db 1981 GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
 Qy 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
 Db 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 Qy 701 LeuGlyGlySerLysTrpGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
 Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
 Qy 721 SerThrMetGlyTrpMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
 Db 2161 AGCACATGGGTACATGCTGCTCATCTGCTCATCTGCTCATCTGCTCATCTGCTCATGATG 2220
 Qy 741 ThrIleAlaTrpThrLysLeuTrpCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
 Db 2221 ACCATTGCTTACCAACAGCTTACTGCAATTTGGACAAGGAGACCTGGAGAATATTGG 2280
 Qy 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780
 Db 2281 GACTGCTCTATGAA-AAAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2339
 Qy 780 sProValAlaPheLeuSerPheSerSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
 Db 2340 CCTGTGGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 2399
 Qy 800 lIleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTrp 820
 Db 2400 AATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2459
 Qy 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTrpVa 840
 Db 2460 CATCTTGTTCATCT 2519
 Qy 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
 Db 2520 CTGGACAGATCAAAACACCCAGCTTGAATGATCAATTAATTAATTAATTAATTAATTAATTA 2579
 Qy 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
 Db 2580 GTCTGTGACTCAACTCAAGCTTGGTAACCTTTACCAGCTCCAGCATCATTGACCT 2639
 Qy 880 uProProSerSerValProSerProAlaTrpProValThrGluSerCysHisLeuSerSe 900
 Db 2640 GCCTCCAGTTCGTCGCCATCACCAGCTTATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2699

QY 900 rValAlaPheValProCysLeu 907
 Db 2700 TGTGGCATTTGTCCCATGTCTC 2721
 RESULT 10
 AF061444
 LOCUS
 DEFINITION
 Homo sapiens G protein-coupled receptor LGR5 (LGR5) mRNA, complete cds.
 ACCESSION AF061444
 VERSION AF061444.1 GI:3885471
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2724)
 AUTHORS Hsu,S.Y., Liang,S.G. and Haueh,A.J.
 TITLE Characterization of two LGR genes homologous to gonadotropin and a thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region
 JOURNAL Mol. Endocrinol. 12 (12), 1830-1845 (1998)
 MEDLINE 99055210
 PUBMED 9849958
 REFERENCE 2 (bases 1 to 2724)
 AUTHORS Hsu,S.Y., Liang,S.G. and Haueh,A.J.W.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-1998) GYN/Ob, Stanford University, 300 Pasteur Dr., Stanford, CA 95305-5317, USA
 FEATURES
 source
 1..2724
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..2724
 /gene="LGR5"
 1..2724
 /gene="LGR5"
 /note="Leucine-rich repeat containing G protein-coupled receptor; similar to glycoprotein hormone receptors; similar to FSH receptor, LH receptor and TSH receptor."
 /codon_start=1
 /product="G protein-coupled receptor LGR5"
 /protein_id="AAC77911.1"
 /db_xref="GI:3885472"
 /translation="MDTSRLGVLSLPVLQLATGSSPRSGVLRGCPTHCHCEPDG RMLLVDCSDLSLSPNSLVFTSIDLSMNNISQLLPNPLPSLHLEELRAGNAL TYIPKGAFTGLYSKLVLMQNNQLRHVPEALQNRSLQSLRDANDHISYVPPSCFSG LHSRLWLDNNALTEIPVQAFRSLSALQAMTALNKIHHIPDYAGNLSVVVLHLH NNRIHSLGKCFDGLHSLBETLDNLYNNLDEFPPTARTLSNLKELGFHNNIRSIPEKA FVGNPSLITIHFDYDNPLOFVGSRSAFOHLPELRTLTLNGASQITERPDLTGTALESIT LTGAQISLSPQTVNQLPNLQVLDLSYNLLEDPFSVCQKLOKIDLRHNEIYEIKVD TFQQLSRLSNLKNATLHPNAPSTLPSLILKDLSSNLSSPFTIGLHGLHLKL TGNHALQSLISENPELVKTEMPYAYQCCAFGVGCENAYKISQNNKNGNSMDDLHK KDAGNFQADRDLEDLDEEDLKLHVSQCSPSPGPKFCEHLLDGLLRIGVYT IAVLATCNALVTSTVFRSPLYISPIKLLIGVIAVNMVLTVSSAFLAGVDAFTFGSFL ARHGAWNGVGVCHVIGFLSIFASSSVFLTLAALERGSVKYSAKPRTKAPFSSLK VIIICALLALTMVAAPVLLGSKYKASPLCLPLPFGEPSTMGVMWALILNLSICLMM TIATYTKLVCNDKGDLEINWCSWVKHALLFTNCLINCPVAFISLNLNITFISP EYIKFLLVYVPLPACLNPLLYLFPNPHKFDLVSRLKQTYVTWYTKSKHSLMSINSDD VEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVPTESCHLSSVAFVPL"
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 2724
 Score: 730.00 Matches: 905
 Percent Similarity: 99.56% Conservative: 0
 Best Local Similarity: 99.56% Mismatches: 2
 Query Match: 80.49% Indels: 4
 DB: 9 Gaps: 0

US-10-751-736-84 (1-907) x AF061444 (1-2724)

Qy	1	Met	Asp	Thr	Ser	Arg	Leu	Gly	Val	Leu	Leu	Ser	Leu	Pro	Val	Leu	Leu	Gln	Leu	Ala	Thr	20						
Db	1	ATG	GAC	ACCT	CCC	GCT	GCG	TGT	GCT	TCT	TGT	GCT	TGT	GCT	TGT	GCT	GCT	GCT	GCT	GCT	GCG	60						
Qy	21	Gly	Gly	Ser	Pro	Arg	Ser	Gly	Val	Leu	Leu	Arg	Gly	Cys	Pro	Thr	His	Cys	His	Cys	40							
Db	61	GGG	GGC	AGC	TCT	CC	CAG	GT	TGT	TGT	GTG	AGG	GGC	TGCC	CC	ACA	CAC	ACT	GTC	ATT	TGC	120						
Qy	41	Glu	Pro	Asp	Gly	Arg	Met	Leu	Leu	Arg	Val	Asp	Cys	Ser	Asp	Leu	Gly	Leu	Ser	Glu	Leu	60						
Db	121	GAG	CCC	GAC	GCG	ACG	AGT	TTG	TCT	CAG	GGT	GG	ACT	TGCT	CCG	ACT	TGG	GGC	TCT	CGG	AGCTG	180						
Qy	61	Pro	Ser	Asn	Leu	Ser	Val	Phe	Thr	Ser	Tyr	Leu	Asp	Leu	Ser	Met	Asn	Asn	Leu	Ser	Gln	80						
Db	181	CC	TTCC	AACT	CAG	CG	CTT	CAC	CTC	TCT	ACT	TAG	ACCT	CAG	TAT	TGA	ACA	CAC	ATC	ATC	ATC	240						
Qy	81	Leu	Leu	Pro	Asn	Pro	Leu	Pro	Ser	Leu	Arg	Phe	Leu	Glu	Glu	Leu	Arg	Leu	Ala	Gly	As	100						
Db	241	CTG	CTCC	GAA	TCC	CTG	CC	CAG	TCT	CCA	-	CTT	CTC	TGG	AGG	AGT	TAC	GTCT	TGG	GGAAA	299							
Qy	100	n	Ala	Leu	Thr	Tyr	Leu	Pro	Lys	Gly	Ala	Phe	Thr	Gly	Leu	Tyr	Ser	Leu	Lys	Val	Leu	Me	120					
Db	300	CGC	TCT	GAC	ATAC	AT	CCC	ACG	GAG	CA	TTC	ACT	TGG	CGC	TTT	TAC	AGT	CTT	TAA	AGT	TCT	359						
Qy	120	T	Leu	Gln	Asn	Asn	Gln	Leu	Arg	His	Val	Pro	Thr	Glu	Ala	Leu	Gln	Asn	Leu	Arg	Ser	Le	140					
Db	360	GCT	CG	AA	TAT	NAT	CAG	CT	TAG	AC	GAC	GAT	CC	ACA	GAG	CTT	TGC	AG	ANT	TTG	GCA	GCCT	419					
Qy	140	u	Gln	Ser	Leu	Arg	Leu	Asp	Ala	Asn	His	Leu	Ser	Tyr	Val	Pro	Ser	Cys	Phe	Ser	G	160						
Db	420	TCA	TCC	TGG	CT	GAT	GT	CA	ACC	ACAT	CAG	CTAT	GT	CCCC	CA	AGCT	TGTT	TC	AGT	TGG	479							
Qy	160	Y	Leu	His	Ser	Leu	Arg	His	Leu	Trp	Leu	Asp	Asp	Ala	Leu	Thr	Glu	Leu	Pro	Val	G	180						
Db	480	CTG	CA	TTC	CC	TG	AG	CA	CC	TGT	GG	CT	TGG	AT	GACA	CA	AT	CG	CTT	TAA	CAAA	TCC	539					
Qy	180	n	Ala	Phe	Arg	Ser	Leu	Ser	Ala	Leu	Gln	Ala	Met	Thr	Leu	Ala	Leu	Asn	Lys	Leu	His	Hi	200					
Db	540	GGC	TTT	TAG	AA	GT	TTT	AT	CGG	CA	TTG	CA	AGC	CA	TG	ACCT	TG	CCCT	TGC	ACAA	AT	TAC	ACCA	599				
Qy	200	s	ile	Pro	Asp	Tyr	Ala	Phe	Gly	Asn	Leu	Ser	Ser	Leu	-	Val	Val	Leu	His	Leu	His	Asn	220					
Db	600	CAT	ACC	AG	ACT	AT	TGCC	TT	TG	AA	ACCT	CT	CC	AG	-	CT	GGT	AGT	TCT	TAC	ATCT	CC	TAACA	658				
Qy	220	s	n	Arg	Leu	His	Ser	Leu	Gly	Lys	Lys	Cys	Phe	Asp	Gly	Leu	His	Ser	Leu	Glu	Thr	Leu	Ala	240				
Db	659	ATA	GAAT	CCA	CT	CC	CTG	GG	AA	GAA	AT	GC	TTT	TGA	TGG	GGC	CT	CCA	CAG	CT	TAG	AG	ACTTT	718				
Qy	240	s	Leu	Asn	Tyr	Asn	Asn	Leu	Asp	Glu	Phe	Pro	Thr	Ala	ile	Arg	Thr	Leu	Ser	Asn	Leu	L	260					
Db	719	ATT	TAA	ATT	TCAA	TAA	CT	TGA	TA	TCC	CA	AT	TCCC	CA	CT	GCA	AT	TAG	GAC	ACT	CT	CA	ACCTTA	778				
Qy	260	Y	s	Glu	Leu	Gly	Phe	His	Ser	Asn	Asn	ile	Arg	Ser	ile	Pro	Glu	Lys	Ala	Phe	Val	Gly	Ala	280				
Db	779	ANG	AACT	AG	AGT	TTT	CAT	AG	CA	CA	AT	TAT	CAG	GTC	GAT	ACCT	TG	AGAA	AGC	AT	TTG	TAG	GCA	838				
Qy	280	s	n	Pro	Ser	Leu	ile	Thr	ile	His	Phe	Tyr	Asp	Asn	Pro	ile	Gln	Phe	Val	Gly	Arg	Ser	Ala	300				
Db	839	ACC	CTT	CT	TAT	TACA	ATA	CAT	TT	TAT	GACA	AT	CC	CA	TCC	AT	TGTT	TGG	GAG	AT	CTG	898						
Qy	300	l	a	Phe	Gln	His	Leu	Pro	Glu	Arg	Thr	Leu	Thr	Leu	Asn	Gly	Val	Ser	Gln	ile	Thr	G	320					
Db	899	CTT	TT	CA	CA	AT	TTT	AC	CT	GAA	CT	TAG	AC	ACT	GAT	CT	GAA	TGG	GGC	CT	CA	AA	TA	ACTG	958			
Qy	320	l	u	Phe	Pro	Asp	Leu	Thr	Gly	Thr	Ala	asn	Leu	Glu	Ser	Leu	Thr	Leu	Thr	Gly	Val	Gln	Ala	340				
Db	959	AA	TTT	CT	GAT	TTT	AA	CT	TGG	AA	CT	GCA	AA	CT	GGA	AGT	CT	GAC	TTT	TAA	CT	TGG	AG	CA	GAGA	1018		
Qy	340	l	e	Ser	Ser	Leu	Pro	Gln	Thr	Val	Cys	asn	Gln	Leu	Pro	asn	Leu	Gln	Val	Leu	Asp	Leu	S	360				
Db	1019	TCT	CAT	CT	TT	CC	T	CA	AA	CC	GT	CT	G	CA	AT	CAG	TT	TAC	CT	TA	AT	CT	CCA	AGT	GCT	GAT	CT	1078

Qy	360	erTyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleA	390
Db	1079	CTTACACCTATTAGAGAATTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATTG	1138
Qy	380	spLeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerL	400
Db	1139	ACCTAAGACATAATGAAATCTACGAAATTAAAGTTGACACTTTTCCAGCAGTGTCTTAGCC	1198
Qy	400	euArgSerLeuAsnLeuAlaTirAsnLysIleAlaIleIleHisProAsnAlaPheSerT	420
Db	1199	TCCGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTATTACCCCAATGCATTTTCCA	1258
Qy	420	hrLeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleT	440
Db	1259	CTTTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCTCTTTCTCTATAA	1318
Qy	440	hrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerIleuI	460
Db	1319	CTGGGTTACATGCTTTAACTCACTTAAATTTAAACAGGAAATCAATGCCTTACAGAGCTTGA	1378
Qy	460	leSerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysC	480
Db	1379	TATCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTAATGCTTACCATGTCT	1438
Qy	480	ysAlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTirAsnLysGlyAspA	500
Db	1439	GTGCAITTTGGAGTGTGAGAAATGCCCTATAAGATTTCTTAATCAATGAATAAAGGTGACA	1498
Qy	500	snSerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluA	520
Db	1499	ACAGCAGTATGGACGACCTTCAATAGAAAGATGCTCGAATGTTTTCAGGCTCAAGATGAAC	1558
Qy	520	rgAspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValG	540
Db	1559	GTGACCTTGGAAGATTTCTCTGCTTGACTTTGAGAGAACCTGAAAAGCCCTTCATTCAGTGC	1618
Qy	540	lnCysSerProSerProGlyProPheLysProCysGluHisLeuLeuLeuAspGlyTyrPleuI	560
Db	1619	AGTGTTCACCTTCCCCAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGATGGCTGGCTGA	1678
Qy	560	leArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrS	580
Db	1679	TCAGAATTGAGTGTGGACCATAGCAGTCTCGCACCTTACTTGTAAATGCTTTTGGTGACTT	1738
Qy	580	erThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleA	600
Db	1739	CAACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCATTTAAACGTTAATTTGGGGTCAATCG	1798
Qy	600	laAlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheT	620
Db	1799	CAGCAGTGAACATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGTGGATGCGGTCA	1858
Qy	620	hrPheGlySerPheAlaArgHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValI	640
Db	1859	CTTTTGGCAGCTTTGCACGACATGCTGCTCGCTGGTGGGAGAAATGGGGTTGGTGGCATGTCA	1918
Qy	640	leGlyPheLeuSerIlePheAlaSerClnSerSerValPheLeuLeuThrLeuAlaAlaL	660
Db	1919	TTGGTTTTTTGTCCATTTTTGCTTTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCC	1978
Qy	660	euGluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerS	680
Db	1979	TGGAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTTGAACAGAAAGCTCCATTTTCTA	2038
Qy	680	erLeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProL	700
Db	2039	GCCTGAAAGTAATCATTTTGTCTGTGCCCTGTGCTGGCTTTGACCATGGCGCAGTTTCCCC	2098
Qy	700	eulLeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluP	720
Db	2099	TGCTGGGTGGGAGCAAGATGGGGCCCTTCCCTCTCTGCTCTGCCCTTTGCCCTTTTGGGGAGC	2158
Qy	720	roSerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetM	740

Db 2159 CACGACACACGCTACATGGTGGCTCTCATCTGCTCAATTCCTTGGCTTCTCATGA 2218
QY 740 etThrileAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleT 760
Db 2219 TGACCATTCCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGGAATATT 2278
QY 760 rPaspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnC 780
Db 2279 GGGACTGCTCTATGGTAAACACATTCGCCCTGTGTCTTCCACCAATGCTAAACT 2338
QY 780 yProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluV 800
Db 2339 GCCCTGTGCTTCTTCT 2398
QY 800 alileLysPheIleLeuLeuValValProLeuProAlaCysLeuLeuAsnProLeuLeuT 820
Db 2399 TAAATTAAGTTTATCTCTCTGTTGGTAGTCCACTTCTGCTGCTCTCAATCCCTTCTCT 2458
QY 820 yIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrV 840
Db 2459 ACATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGGCTTGAGAAAGCAACCTAGC 2518
QY 840 alTnThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysG 860
Db 2519 TCTGGACAAGATCAAAACACCCCAAGCTTGTATGCAATTAACCTGTGATGATGTCGAAAAC 2578
QY 860 InSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspL 880
Db 2579 AGTCTGTGACTCACTCAAGCTTGTGTAACCTTTACAGCTCCAGCATCACTATGACC 2638
QY 880 euProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerS 900
Db 2639 TGCCCTCCAGTTCCGTCGCTCAACAGCTTATCCAGTACTGAGAGTGCATCTTCTCT 2698
QY 900 exValAlaPheValProCysLeu 907
Db 2699 CTGTGGCATTTGTCCATGCTCTC 2721

BD135244 2082 bp DNA linear PAT 18-SEP-2002
LOCUS Novel mammalian G protein-coupled receptor having extracellular
DEFINITION leucine-rich repeating domain.
ACCESSION BD135244
VERSION BD135244.1 GI:23230189
KEYWORDS JP 2002507406-A/2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.
TITLE Novel mammalian G protein-coupled receptor having extracellular
JOURNAL leucine-rich repeating domain
Patent: JP 2002507406-A 2 12-MAR-2002;
THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,
AKZO NOBEL NV

COMMENT
OS Homo sapiens (human)
PN JP 2002507406-A/2
PD 12-MAR-2002
PF 25-MAR-1999 JP 2000537903
PR 26-MAR-1998 US 60/079501
PI ARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES
PI VAN DER SPEK
PC C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,
PC C12N1/21,
PC C12N5/10,C12P21/08,C12N15/00,C12N5/00
CC Novel mammalian G protein-coupled receptor having CC
extracellular
CC leucine-rich repeating domain
FH Key
FT Location/Qualifiers
FT source 1. .2082

FT Location/Qualifiers
FEATURES Source
1. .2082
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2082
Score: 606.00 Matches: 692
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 66.81% Indels: 2
DB: Gaps: 0

US-10-751-736-84 (1-907) x BD135244 (1-2082)

QY 215 LeuHisLeuHisAsnAsnArgIleHisSerLeuGlyLysCysPheAspGlyLeuHis 234
Db 1 CTACATCTCCATAACAATAGATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCAC 60
QY 235 SerLeuGluThrLeuAspLeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArg 254
Db 61 AGCCTAGAGACTTTAGATTTAAATTACAAATTAACCTTGATGAATCCCCACTGCANTAGG 120
QY 255 ThrLeuSerAsnLeuLysGluLeuGlyPheHisSerAsnAsnIleArgSerIleProGlu 274
Db 121 ACATCTCCCAACTTAAAGGAACCTAGGATTTCTAGCAACAATATCAGGTCGATACCTGAG 180
QY 275 LysAlaPheValGlyAsnProSerLeuIleThrIleHisPheTyrAspAsnProIleGln 294
Db 181 AAAGCATTTTGTAGGCAACCTCTCTTATTACAAATACATTTCTATGACAAATCCCATCAA 240
QY 295 PheValGlyArgSerAlaPheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGly 314
Db 241 TTGTGGGAGATCTGCTTTTCAACATTTACCTGAACTAGAACACTGACTCTGANTGGT 300
QY 315 AlaSerGlnIleThrGluPheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThr 334
Db 301 GCCTCACAATAAATGAATTTCTCTGATTTAACTGGAACCTGCAACCTGGAGAGCTGACT 360
QY 335 LeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeu 354
Db 361 TTAACCTGGAGACACAGATCTCATCTCTCTCTCAACCGCTCTGCAATCAGTTACCTAATCTC 420
QY 355 GlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPheSerValCysGln 374
Db 421 CAAGTCTAGATCTGTCTTACAACTATTAGAAAGATTTACCCAGATTTTTCAGTCTGCCAA 480
QY 375 LysLeuGlnLysIleAspLeuArgHisAsnGluIleTyrGluIleLysValAspThrPhe 394
Db 481 AGCTTCAGAAATTTGACCTTAGACATATGAATGAATCTAGAAATTAAGATTGACACTTTC 540
QY 395 GlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHis 414
Db 541 CAGCAGTTCTTAGCTCCGATCCGATCGCTGAATTTGGCTTGGAAACAAATAATGCTATTATTCAC 600
QY 415 ProAsnAlaPheSerThrLeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeu 434
Db 601 CCCAATGCAATTTTCCACTTTGCCATCCCTAAATAAGCTGGACCTATCGTCCAACTCCCTG 660
QY 435 SerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHis 454
Db 661 TGTCTTTTCTTAACTAGCTGGGTACATGTTTAACTCACTTAAATAATACAGGAATCAT 720
QY 455 AlaLeuGlnSerLeu-IleSerSerGluAsnPheProGluLeuLysValIleGluMetPr 474
Db 721 GCCTTACAGAG-CTGGATATCATCTGAAAACCTTCCAGAACTTCCAGAACTNATAGAAATGCC 779
QY 474 cTyrAlaTyrGlnCysCysAlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGln 494
Db 780 TTATGCTTACCAGTGTCTGTGCATTTTGGAGTGTGTGAGAAATGCCATATAAGATTTCTAATCA 839

Qy	494	nTrpAsnLysGlyAspAsnSerSerMetAspAspLeuHisLysLysAspAlaGlyMetPh	514
Db	840	ATGGAAATAAGGTGACAAACAGCAGTATGGAGCAGCTTCATAAGAAAGATGCTGGATGTT	899
Qy	514	eGlnAlaGlnAspGluArgAspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLys	534
Db	900	TCAGGCTCAAGATGAACGTGACCTTGAAGATTTCCTGCTTGACCTTTGAGGAAGACCTGAA	959
Qy	534	sAlaLeuHisSerValGlnCysSerProSerProGlyProPheLysProCysGluHisLe	554
Db	960	AGCCCTTCATTACGTGAGTGTTCACCTTCCCAGGCCCTTCAAACCCCTGTGAACACCT	1019
Qy	554	uLeuAspGlyTrpLeuLeuArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCys	574
Db	1020	GCTTGATGGCTGGCTGATCATCAGATTGAGATGTGACCATAGCAGTCTCGCAGCTTACTTG	1079
Qy	574	sAsnAlaLeuValThrSerThrValPheArgSerProLeuTyrlleSerProIleLysLe	594
Db	1080	TAATGCTTTGGTGACTTCAACAGATTTTCAGATCCCTCTGTACATTTCCCCCAATAAACT	1139
Qy	594	uLeuIleGlyValIleAlaValAsnMetLeuThrGlyValSerSerAlaValLeuAla	614
Db	1140	GTTAAATTTGGGTGCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGTGGC	1199
Qy	614	aGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyValaTrpTrpGluAsnGly	634
Db	1200	TGGTGTGGATGCGTTTCATCTTTTGGCAGCTTTGGACAGCATGGTGCCCTGGTGGGAGAAATGG	1259
Qy	634	yValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGluSerSerValPheLe	654
Db	1260	GTTTGGTTGCCATGTTCATTTGCTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCCT	1319
Qy	654	uLeuThrLeuAlaAlaLeuGluArgGlyPheSerValLysTrsSerAlaLysPheGluTh	674
Db	1320	GCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAAATATATCTGCAGAAATTTGAAAC	1379
Qy	674	rLysAlaProPheSerSerLeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuTh	694
Db	1380	GAAGCTCCATTTCTAGCTGGAAGTAATCATTTGCTCTGTGCCCTGTGCGCTTGAC	1439
Qy	694	rMetAlaAlaValProLeuLeuGlyLysSerLysTrsGlyValaSerProLeuCysLeuPr	714
Db	1440	CATGGCCGCAGTTTCCCTGTCTGGTGGCAGCAGTATGGCGCTTCCCTCTCTGCGCTGCC	1499
Qy	714	oLeuProPheGlyGluProSerThrMetGlyTrsMetValaLeuLeuLeuLeuAenSe	734
Db	1500	TTTGCTTTTGGGAGCCAGCACCATGGGGTACATGGTCGCTCTCATCTTGCTCAATTC	1559
Qy	734	rLeuCysPheLeuMetMetThrIleAlaTyThrLysLeuTyrcysAsnLeuAspLysGly	754
Db	1560	CCTTTGTCTCCTCATGATGACCATTTGGCTACACCAAGCTCTACTGCAATTTGGACAAGG	1619
Qy	754	yAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAlaLeuLeuLeuPheTh	774
Db	1620	AGACTGGAGAAATTTTGGGACTGCTCTATGGTAAACACACATTGCCCTGTGTCTTTCAC	1679
Qy	774	rAsnCysIleLeuAenCysProValAlaPheLeuSerPheSerSerLeuIleAenLeuTh	794
Db	1680	CAACTGCATCTAAACTGCCCTGTGGCTTTCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCT	1739
Qy	794	rPheIleSerProGluValIleLysPheIleLeuLeuValValProLeuProAlaCys	814
Db	1740	ATTTATCAGTCTCGAAGTAAATTAAGATTTTATCTCTTCTGGTGGTAGTCCACCTTCTGTGATG	1799
Qy	814	sLeuAsnProLeuLeuTyrlleLeuPheAsnProHisPheLysGluAspLeuValSerLe	834
Db	1800	TCTCAATCCCCCTCTCTACATCTTGTTCATCTCTTCAATCTCTTCAAGGAGATCTGGGAGCCT	1859
Qy	834	uArgLysGlnThrTyrlleValTrpThrArgSerLysHisProSerLeuMetSerIleAenSe	854
Db	1860	GAGAAAGCAACCTACGCTCTGCAACAGATCAAAACACCAAGCTTGATGTCAATTAATCT	1919

Qy	854	rAspAspValGlulylveGlnSerCysAspSerThrGlnAlaLeuValThrPheThrSerSe	874
Db	1920	TGATGATGTCGAAAAACAGTCTGTGACTCAACTCAAGCCCTTGTTACACCTTTACAGCTC	1979
Qy	874	rSerIleThrThyAspLeuProProSerSerValProSerProAlaTyProValThrGl	894
Db	1980	CAGCATCACATTATGACCTGCCTCCCACTCCGTCGACCATCCAGCTTATCCAGTGACTGA	2039
Qy	894	uSerCysHisLeuSerSerValAlaPheValProCysLeu	907
Db	2040	GAGTCGCACACTTTCCTCTGTGGCAATTGTCACCATGCTC	2079
RESULT 12			
AC078860/c			
LOCUS		145165 bp	DNA linear PRI 23-JAN-2003
DEFINITION		Homo sapiens 12 BAC RP11-186F10 (Roswell Park Cancer Institute	
ACCESSION		Human BAC Library) complete sequence.	
VERSION		AC078860	
KEYWORDS		HTG.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 145165)	
		Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C.,	
		Alsbrooks,S.L., Amaratingue,H.C., Are,J.R., Banks,T., Barbara,J.,	
		Benton,J., Bimonge,K., Blankenburg,K., Bonnin,D., Bouck,J.,	
		Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,	
		Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,	
		Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,	
		Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,	
		Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,	
		Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,	
		Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,	
		Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,	
		Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,	
		Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,	
		Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,	
		Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,	
		Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O.,	
		Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsai,F.,	
		Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,	
		Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,	
		Karlsson,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,	
		Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,	
		Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,	
		Loulesged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,	
		Ma,J.J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,	
		Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,	
		Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,	
		Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,	
		Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,	
		Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S.,	
		Ogoh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,	
		Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,	
		Quiles,M., Ren,J., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,	
		Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,	
		Shooshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,	
		Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,	
		Therisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,	
		Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,	
		Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,	
		Washington,C., Watlington,S., Williams,G., Williamson,A.,	
		Wlaczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Zhou,J.,	
		Zorilla,S., Kucherlapati,R. and Gibbs,R.	
		Direct Submission	
		Unpublished	
		2 (bases 1 to 145165)	
JOURNAL		Worley,K.C.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (07-AUG-2000) Human Genome Sequencing Center, Department	
TITLE		of Molecular and Human Genetics. Baylor College of Medicine. One	
JOURNAL			

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (31-MAR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (01-APR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (23-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 2001 this sequence version replaced gi:13324671.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

Contig length: 145169
Phrap values in estimate: 144455
Average error rate (BCM-Phrap estimate): 5.28339e-05
Fraction of Phrap values less than 40 : 0.0137759
Number of consensus changing edits: 19
Number of N's in consensus : 0

Summary Statistics

Contig length: 145169
Phrap values in estimate: 144455
Average error rate (BCM-Phrap estimate): 5.28339e-05
Fraction of Phrap values less than 40 : 0.0137759
Number of consensus changing edits: 19
Number of N's in consensus : 0

Consensus changing edits

Position	Original+Context	Edited+Context
6882	agatgtgat(a)gctgtgctt	agatgtgat(g)gctgtgctt
12828	tttgggtca(n)tcgaacctcc	tttgggtca(c)tcgaacctcc
13151	aaacacacat(n)caatgataa	aaacacacat(a)caatgataa
28957	ttgtattgca(n)atggcatctt	ttgtattgca(a)atggcatctt
37139	atcgttattg(n)tgtaatttt	atcgttattg(t)tgtaatttt
58030	tttttatt(n)tatagttaa	tttttatt(t)tatagttaa
58290	tcacgacct(a)tcagtccct	tcacgacct(a)tcagtccct
64729	ctatctctac(n)aaagaaga	ctatctctac(g)aaagaaga
92472	aatggcttt(g)aatgggccc	aatggcttt(g)aatgggccc
92662	cttagactgt(n)aatgcagga	cttagactgt(g)aatgcagga
103773	gctaattttt(n)catctttgt	gctaattttt(a)catctttgt
114056	cacatgcatt(n)ggatttggc	cacatgcatt(a)ggatttggc
117534	tcaagcttgt(n)attcagggt	tcaagcttgt(a)attcagggt
117581	cagatcatct(n)agggaatgct	cagatcatct(g)agggaatgct
125485	gcttacacta(n)aaatgaacc	gcttacacta(g)aaatgaacc
133723	agcattggtgt(n)ncaactaat	agcattggtgt(c)ggcacttaac
133724	gcattggtgt(n)ncaactaata	gcattggtgt(g)gcacttaata
133725	catggtgt(n)ncaactaataa	catggtgt(g)gcacttaataa
140801	aaatatgtc(n)taataaagt	aaatatgtc(c)taataaagt

Distribution of Quality < 40 Bases

	5	10	15	20	25	30	35	40
1000								
900								*
800								*
700								*
600								*
500								*
400	*						*	*
300	*	*				*	*	*
200	*	*	*		*	*	*	*
100	*	*	*	*	*	*	*	*
0	*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 qxf.

FEATURES
Source

Location/Qualifiers
1..145165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="12"
/clone="RP11-186F10"
misc_feature
1..2004
/note="overlaps bases 52259..54262 of clone AC073612"
/function="clone overlap"
repeat_region
complement(6..223)
/rpt_family="LIME1"
repeat_region
complement(404..698)
/rpt_family="LIM4"
repeat_region
complement(833..1077)
/rpt_family="LIM4"
repeat_region
complement(1228..1482)
/rpt_family="LIM4"
repeat_region
complement(1740..2021)
/rpt_family="AluJB"
repeat_region
complement(2098..2151)
/rpt_family="LIMEC"
repeat_region
complement(2152..2449)

Alignment Scores:
Pred. No.: 0 Length: 145165
Score: 362.00 Matches: 362
Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.91% Indels: 0
DB: 9 Gaps: 0

US-10-751-736-84 (1-907) x AC078860 (1-145165)

QY 546 GlyProPheLeuProCysGluHisLeuLeuAspGlyTrpLeuileAArgileGlyValTrp 565
DB 72537 GGGCCCTCAAAACCTGTGAACACCTGCTGTGAAGTGGCTGCTGATCAGAAATGGAGTGG 72478

QY 566 ThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585
DB 72477 ACCATAGCAGTTCTGGCATTACTTGTAAATGCTTTGTGTGACATTCACAGTTTTTCAGATCC 72418

QY 586 ProLeuTyrlleSerProIleLysLeuLeuileGlyValIleAlaAlaValAsnMetLeu 605
DB 72417 CCTCTGTATCATTTCCCCCATTAACCTGTAATTTGGGGTCATCGCAGCAGTGAACATGCTC 72358

QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625
DB 72357 ACGGGAGTCTCCAGTGCCTGCTGGCTGTGGATGCGTTTCACATTTTGGCGAGCTTTGCA 72298

QY 626 ArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIle 645
DB 72297 CGACATGGTGGCTGTGGGAGAAATGGGGTGGTGGTCCATGTCATTTTGTCCATT 72238

QY 646 PheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSer 665
DB 72237 TTTGGCTTCAGATCATCTGTTTTCCTGCTTACTCTGGCAGCCTCGAGCGTGGGTCTCT 72178

QY 666 ValLysTyrlleSerAlaValPheGluThrLysAlaPheSerSerLeuLysValIleIle 685
DB 72177 GTGAATATTTGCAAAATTTGAAACGAAAGTCCATTTTCAGCTGAAAGTAATCATTT 72118

QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyLysSerLys 705
DB 72117 TTGCTCTGGCCCTGTGGCCTTGACCATGGCCGAGTTCCTCTGCTGGTGGCAGCAAG 72058

QY 706 TyrlleAlaSerProLeuLeuProLeuProPheGlyGluProSerThrMetGlyTyrlle 725
DB 72057 TATGGCGCTCCCTCTCTGCTGCTTTGGCTTTGGCGAGCCACACCATGGGCTAC 71998

QY 726 MetValAlaLeuLeuLeuLeuLeuSerLeuLysPheLeuMetMetThrIleAlaTyrlle 745
DB 71997 ATGGTGGCTCTCATCTTGCTCAATTTCCCTTTCCTCATGATGACCATTCCTACAC 71938

QY 746 LysLeuTyrlleCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetVal 765
DB 71937 AAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAATATTTGGAGCTGCTCTATGTA 71878

QY 766 LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeu 785
DB 71877 AAACACATTTGGCTGTCTTCCACCACTGCATCTAACTGACCTGCTGCTGCTTTCTTG 71818

QY 786 SerPheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeu 805
DB 71817 TCCTTCTCTCTTTAATAAACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCCTT 71758

QY 806 LeuValValProLeuProAlaCysLeuAsnProLeuLeuTyrlleLeuPheAsnPro 825
DB 71757 CTGGTGGTAGTCCCACTTCTGCTGCTCAATTCCTTCTACATCTTGTTCATCTCT 71698

QY 826 HisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrlleTrpThrArgSerLys 845
DB 71697 CACTTTAAGGAGGATCTGGTAGCCTCGAGAAAGCAACCTACCTCTGGACAAGATCAAAA 71638

QY 846 HisProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThr 865
DB 71637 CACCAAGCTTGTATGCAATTAACCTCTGATGATGTCGAAACACAGTCTGTGACTCAACT 71578

QY 866 GlnAlaLeuValThrPheThrSerSerSerIleThrTyrlleAspLeuProProSerSerVal 885
DB 71577 CAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCTCCAGTTCCGCTG 71518
```

```
QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnLeuSerGln 80
Db 278 CTTTCCAACCTCAGCGCTTTACCTCCTAGAGCTCAGTATGAACAACATCAGTCAG 337
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
Db 338 CTGCTCCCAATCCCTCGCCAGCTCCGCTTCTCGAGAGGATGATGCTCTTGGGGAAC 397
QY 101 AlaLeuThrTyrLeuProGlyAlaPheThrGlyLeuTyrSerLeuLeuValLeuMet 120
Db 398 GCTCTGACATACATCCCAAGGAGGAGCATTCACCTGGCTTTACAGCTCTTAAGTCTTATG 457
QY 121 LeuGlnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 458 CTGCAGATAATCAGCTAAGACACGTAACCCACAGAAGCTCTGCAGAAATTTGGCAAGCCTT 517
QY 141 GlnSerLeuArgLeuAspAlaLeuHisLeuSerTyrValProSerCysPheSerGly 160
Db 518 CAATCCCTCGGCTGGATGCTAACCAATCAGCTATGTGCCCAAGCTTTTCAGTGGC 577
QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluLeuProValGln 180
Db 578 CTGCATTCCTTGAGGACCTGTGGCTGGATGACATGGCTTAACAGAAATCCCGTCCAG 637
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysLeuHis 200
Db 638 GCTTTTAGAAGTTATTCGGCATTCGAAGCATGACCTTGCCCTGAACAAATACACCAC 697
QY 201 IleProAsnTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn 220
Db 698 ATACCAAGATGATGCCCTTTGGAAACCTCTCAGCTTGGTAGTCTACATCTCCATAACAAT 757
QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 758 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTTAGAGCTTAGAT 817
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 818 TTAATATTAACAATCCTGATGAATTCCTCCACTGCAATAGGACACTCTCCAACCTTAA 877
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 878 GAACATAGGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGCAAC 937
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 938 CTTCTCTTTATACATACATTTCTATGACATCCCATCCATTTGTTGGAGATCTGCT 997
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 998 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTCAATGGTGCCTCACAATTAACCTGAA 1057
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1058 TTTCTCTGATTTAACTGGAACCTGCAACCTAGAGAGTCTGACTTTAACTGAGACACAGATC 1117
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeu 357
Db 1118 TCATCTCTTCTCAAAACCGTCTGCAATCATGTTACCTTAATCTCCAGTGGCTG 1168

RESULT 14
LOCUS E42915
DEFINITION Novel gonadotropin receptor.
ACCESSION E42915
VERSION E42915.1 GI:18627551
KEYWORDS JP 2000125884-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 606)
AUTHORS Heikopu, Y.C. and Supekku, P.J.V.D.
```

```
TITLE Novel gonadotropin receptor
JOURNAL Patent: JP 2000125884-A 3 09-MAY-2000;
COMMENT AKZO NOBEL NV
OS Unidentified
PN JP 2000125884-A/3
PD 09-MAY-2000
PF 05-FEB-1998 JP 1999029120
PF 06-FEB-1998 EP 98200357.6, 27-JUL-1998 EP 98202519.9, PR
24-SEP-1998 EP 98203213.8
PI YUDEISU CHRISTINA HEIKOPU, PETORIYUSU JOHANNES VAN DEL SUPEKKU
PC C12N15/09, A61K38/00, A61P15/08, A61P43/00, C07K14/72, C12N1/15, PC
C12N1/19
PC C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, PC
G01N33/566,
PC C12N15/00, A61K37/02, C12N5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..606 /organism='Unidentified'.
FEATURES
source
Location/Qualifiers
1..606
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.: 7,04e-196 Length: 606
Score: 202.00 Matches: 202
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.27% Indels: 0
DB: 6 Gaps: 0
US-10-751-736-84 (1-907) x E42915 (1-606)
QY 562 IleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThr 581
Db 1 ATTTGGAGTGTGGACCATAGCAGTCTGGCACCTACTTGTAAATGCTTTGGTGAAGCTTCAACA 60
QY 582 ValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 601
Db 61 GTTTTTCAGATCCCTCTGTACATTTCCCTCCATTTAAACTGTAAATGGGTGTCATCGCAGCA 120
QY 602 ValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPhe 621
Db 121 GTGAACATGCTCAGGGAGTCTCCAGTCCGCTGGCTGGTGGATGGGTTCATCTTTT 180
QY 622 GlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGly 641
Db 181 GGCAGCTTTTGCACGACATGGTGCCTGGTGGAGAAATGGGGTGGTTCATGTCATGTTGGT 240
QY 642 PheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeuGlu 661
Db 241 TTTTGTGTCATTTTGTCTTTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAG 300
QY 662 ArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaPheProPheSerSerLeu 681
Db 301 CGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCATTTTCTAGCCTG 360
QY 682 LysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeu 701
Db 361 AAAATAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 702 GlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSer 721
Db 421 GGTGGCAGCAGATATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 722 ThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMetThr 741
Db 481 ACCATGGGCTACATGCTGCTCTCATCTTGTCAATTCCTTGTCTTCTCATGATGACC 540
```


Db 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGTCTGAGGGGCTGCCACACACTGTCTATGTC 120
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 121 GAGCCCGACGGCAGAGATGTGCTCAGGGTGGACTGCTCCGAGCTGGGGCTCTCGAGGCTG 180
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 181 CCTTCCACCTCAGGGCTTCACTCTCTACCTAGACCTCAGTATGAAACACACATCAGTCAG 240
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuLeuLeuAlaGlyAsn 100
Db 241 CTGCTCCCGAAATCCCTCGCCAGTCTCCGCTTCTCGAGGAGTACGCTCTTGGCGGAAAC 300
Qy 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTTATG 360
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 361 CTGCAGATATATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGGCAGACCTT 420
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Db 481 CTGCAATCCCTGAGGACCTGTGGCTGGATGCAATGCGTTAAACAGAAATCCCGCTCCAG 540
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db 541 GCTTTTAGAAGTTATCGGCATGTGAGCCATGACCTTGGCCCTGAAACAAATACACAC 600
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuHisLeuHisAsnAsn 220
Db 601 ATACAGGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 661 AGAATCCATCCCTCGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 721 TTAATTTACAAATCTTGATGAATTTCCCACTGCATTTAGACACTCTCCAAACCTTAAA 780
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 781 GAACCTAGGATTTTCATAGCAACAATATCAGGTGCTACCTTGAGAAAGCAATTTGTAGGCAAC 840
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 841 CCTCTCTTTATTAACAATAATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 901 TTTCAACATTTACCTGGAACATAGAACACTGACTCTGAATGGTGGCTTCACAAATAACTGAA 960
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 961 TTTCTGATTTAACTGGAACTGCMAACCTGGAGAGCTGACTTTAACTGGAGCACAGATC 1020
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1081 TACAACTTATAGAGATTTACCCAGTTTTTTCAGTCTGCGAAAGCTTCAGAAATTTGAC 1140
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400

Db 1141 CTAAGACATAATAAGAAATCTACGAAATTAAGTTGACATCTTCCAGCAGTTGCTTAGCCTC 1200
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAAATAAGCTGGACCTATCGTCCAACCTCTCTGCTCTTTTCTTAATACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAACTTTCAGAACTCAGGTTATAGAAATGCCTTATGCTTACCAGTGTCTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnIleProAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Qy 501 SerSerMetAspAspLeuHisLysIlyAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAAGATTTCTGCTTGCCTTGCAGAGACCTTGAAGGCCCTTCATTCAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleLeu 560
Db 1621 TGTTTCCACTTCCCCAGGCCCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATTTGGAGTGTGGACCATAGCAGATTTCTGACACTTACTTGTAAATGCTTTGGTGAATTC 1740
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAATCTGTTAATTTGGGGTCAATCGCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTAGAACATGCTCACGGAGATCTCCAGTGCCTGCTGGCTGTGTGGATGCGTTCCTACT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGGAATGGGGTTGGTTGCCATGTCAAT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCACTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGGTCTCTGTGAATATTTTCGAAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGGCTGTCGCTTGGACCATGGCGCAGTTCCCTG 2100
Qy 701 LeuGlyGlySerLysTyrGlyValSerProLeuCysLeuProLeuPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGATGGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGAGCCCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGCTACATGTCGCTCTCATCTTGCTCAATTCCTTTGCTTCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTTGCTTACACCAGCTCTACTGCAATTTTGGCAAGGGGAGACCTGTGAGAATATTTGG 2280

Qy	761	AspCysSerMetValIysHisIleAlaLeuLeuLeuPheThrAenCysIleLeuAenCys	780
Dd	2281	GACTGCTCTATGGTAAACAACATTGCCCTGTGTGCTTCACCAACTGCATCTCAAACCTGC	2340
Qy	781	ProValAlaPheLeuSerPheSerSerLeulleasnLeuThrPheIleSerProGluVal	800
Dd	2341	CCGTGGCTTTCTTGCTCTTCTCTCTTTTAATAAACCTTACATTTATCATGCTCTGAAGTA	2400
Qy	801	IleIysPheIleLeuLeuValValProLeuProAlaCysLeuAenProLeuLeuTyr	820
Dd	2401	ATTAAAGTTTATCTCTCGTGGTAGTCCCACCTTCCTGCATGTCTCAATCCCCCTTCTCTAC	2460
Qy	821	IleLeuPheAenProHisPheLysGluaspLeuValSerLeuaglysglnThrTyrVal	840
Dd	2461	ATCTGTGTCAATCTCTCACTTTAAGAGGAGTCTGGTGAGCCTGAGAAGCAAACTCAGTC	2520
Qy	841	TrpThrArgSerIysHisProSerLeuMetSerIleAasnSerAspAspValGluLysGln	860
Dd	2521	TGGACAAGATCAAAACACCACCAAGCTTCATGTCAATTAACTCTGATGATGTCGAAAAACAG	2580
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Dd	2581	TCCTGTGACTCACTCAAGCCTTGTTAACTTTTACCAGCTCCAGACATCACTTATGACCTG	2640
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Dd	2641	CCTCCAGTTCCTGGGCCATCACCAAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2700
Qy	901	ValAlaPheValProCysLeu	907
Dd	2701	GTGGCATTTGTCCCATGTCTC	2721

RESULT 2

US-09-976-594-201
Sequence 201, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 201
LENGTH: 4570
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 205542.2
NAME/KEY: unsure
LOCATION: 3900, 3919, 3934
OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-201

No colour

Alignment Scores:	
Pred. No.:	0
Score:	907.00
Length:	4570
Matches:	907
Conservative:	0
Percent Similarity:	100.00%
Mismatches:	0
Best Local Similarity:	100.00%
Query Match:	100.00%
Indels:	0
Gaps:	0
DB:	4

US-10-751-736-84 (1-907) x US-09-976-594-201 (1-4570)

Qy	1	Met	Asp	Thr	Ser	Arg	Leu	Gly	Val	Leu	Leu	Ser	Leu	Pro	Val	Leu	Gln	Leu	Ala	Thr	20
Db	262	ATG	GA	CAC	CTC	CCG	GT	CGG	TG	TCT	CT	CTT	GGCT	GT	GT	CTG	CTG	CAG	TGG	CG	321

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB 121 GAGCCGACGCGAGGATGTGCTCAGGGTGAGCTGCTCCGACCTCGGGCTCTCGGAGCTG 180
QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB 181 CTTTCCAACTCAGGCTCTTACCTCTTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
DB 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTACGCTCTGCGGGAAAC 300
QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
DB 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTTATG 360
QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB 361 CTGCGAGAATAATCAGCTAAGACACGTACCACAGAGCTCTCGCAATTTTGGGAAGCCTT 420
QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB 421 CAATCCCTCGCTCTGGATGCTTAACACATCAGTATGTGCCCCCAAGCTGTTTCAGTGGC 480
QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
DB 481 CTGCATTCCTGAGGACCTGTGGCTGGATGACATGCGTTAAACAGAAATCCCGCTCCAG 540
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
DB 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAACAAATATACCCAC 600
QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
DB 601 ATACAGACTATGCTTTGGAACTCTCCAGCTTGTAGTCTTACATCTCCATTAACAAT 660
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 661 AGAATCCACTCCTCGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 721 TTAATATCAATAAACCTTTGATGAATTCCTCCACTGCAATTTAGGACACTCTCCAACTTAA 780
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 781 GAACTAGGATTTATAGCAACAAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 841 CTTCTCTCTATTACAAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCAAAATAACTGAA 960
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 961 TTTCTCTGATTTTAACTTGGAACTGCAAACTGGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTTCTCAAACTGCTGCAATCAGTTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACTTATTAGAAAGATTTTACCCAGTTTTTTCAGTCTGCGCAAAAGCTTTCAGAAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAAGACATAATGAATATCTACGAAATTAAGTTTGAACATTTTCCAGAGTGTCTTAGCCTC 1200

QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAAATTTGGCTTGGAAACAAAATTGTATATTATTCACCCCAATGCATTTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTTCTGTCTCTTTCTCTATACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCATTATGCTTTACAGTGTCTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGAGTGTGTGAGAAATGCCCTATAGATTTCTATCAATGAATANAGGTGACAAAC 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCAAGGCCCTTCAACCTGTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATAGCAGCTTCTGGCACCTTACTTGTATGCTTTTGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCATTTAAACTGTTTAAATGGGGTCACTCGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTCCCGTGTGGCTGGTGTGGATGGCTTCACT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGSCACCTTTGCACGACATGGTCCCTGGTGGGAGAAATGGGGTGGTGGCTGCTGCTATT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTTTGTCCATTTTGTCTTCAATCATCTGTTTTTCTGCTTACTCTTGGCAGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCTGGGGTCTCTGTGAAATATTTCTGCAAAATTTTGAACAGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2041 CTGAAAGTAAATCATTTTGTCTGTGGCTGTGGCTTGGACCATGGCCGCGAGTTCCTCCCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuLysLeuProLeuProPheGlyGluPro 720
DB 2101 CTGGTGGCAGCAAGATATGGGCCCTCCCTCTCTGCTGCTTGGCTTGGCTTGGGGAGCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuAsnSerLeuLysPheLeuMetMet 740
DB 2161 AGCACATGGGTACATGGTCTCATCTGTCTGCTCAATTCCTTGTCTTCTCTCAATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
DB 2221 ACCATTGCTACCAACAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTGG 2280
QY 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2484)
US-09-799-451-723

Alignment Scores:
Pred. No.: 2,52e-07 Length: 2988
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-799-451-723 (1-2988)

Qy 285 ThrIleHisPheTyraAspAsnProIleGlnPheValGlyArgSerAlaPheGln 302
Db 436 ACATACACTTTTATGATACCAATCCAGTTTGTGGAGATCGGCATTCACG 489

RESULT 6
US-09-495-050A-214
; Sequence 214, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1
US-09-495-050A-214

Alignment Scores:
Pred. No.: 0.0188 Length: 2612
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-495-050A-214 (1-2612)

Qy 208 AsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 266 AACCTTCAAGCGCTGGTAGTCTTCGCATCTTCATACAAAT 304

RESULT 7
US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
```

```
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-866-757-1

Alignment Scores:
Pred. No.: 0.031 Length: 4203
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-84 (1-907) x US-08-866-757-1 (1-4203)

Qy 208 AsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 158 AACCTTCAAGCGCTGGTAGTCTTCGCATCTTCATACAAAT 196

RESULT 8
US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4203
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-153-593-1

Alignment Scores:
Pred. No.: 0.031 Length: 4203
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-84 (1-907) x US-09-153-593-1 (1-4203)

Qy 208 AsnLeuSerLeuValValLeuHisLeuHisAsnAsn 220
Db 158 AACCTTTCAGCCTGTAGTCTGCACTTTCATACAAAT 196

RESULT 9

US-09-397-787-240
; Sequence 240, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 240
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-397-787-240

Alignment Scores:
Pred. No.: 0.283 Length: 453
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-84 (1-907) x US-09-397-787-240 (1-453)

Qy 810 ProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 120 CCCCTGCTGCTGCTGCTCAACCACTGCTGTAC 152

RESULT 10

US-09-556-916-25
; Sequence 25, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-556-916-25

Alignment Scores:
Pred. No.: 128 Length: 2019
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-556-916-25 (1-2019)

Qy 708 AlaSerProLeuCysLeuProLeuPro 716
Db 1892 GCCTCCCTTTGTGCTGCTTTGCCA 1918

RESULT 11

US-09-556-916-27
; Sequence 27, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-556-916-27

Alignment Scores:
Pred. No.: 128 Length: 2019
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-556-916-27 (1-2019)

Qy 708 AlaSerProLeuCysLeuProLeuPro 716
Db 1892 GCCTCCCTTTGTGCTGCTTTGCCA 1918

RESULT 12

US-09-949-016-5657
; Sequence 5657, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5657
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5657

Alignment Scores:
Pred. No.: 172 Length: 2680
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-949-016-5657 (1-2680)

QY 95 LeuArgLeuAlaGlyAsnAlaLeuThr 103

DB 2324 CTGAGACTGGCTGGAATGCTCTGACT 2350

RESULT 13

US-09-799-451-422

; Sequence 422, Application US/09799451

; Patent No. 6783969

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

; APPLICANT: Xue, Aidong J.

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Ma, Yungqing

; APPLICANT: Yamazaki, Victoria

; APPLICANT: Chen, Rui-hong

; APPLICANT: Wang, Zhiwei

; APPLICANT: Wang, Dunrui

; APPLICANT: Yang, Yonghong

; APPLICANT: Wehrman, Tom

; APPLICANT: Ghosh, Reena

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6783969el Nucleic Acids and

; FILE REFERENCE: 803

; CURRENT APPLICATION NUMBER: US/09/799,451

; CURRENT FILING DATE: 2001-03-05

; NUMBER OF SEQ ID NOS: 948

; SOFTWARE: pt_FL_genes Version 2.0

; SEQ ID NO 422

; LENGTH: 2820

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (265)..(2577)

US-09-799-451-422

Alignment Scores:
Pred. No.: 181 Length: 2820
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-799-451-422 (1-2820)

QY 708 AlaSerProLeuCysLeuProLeuPro 716

DB 2156 GCCTCCCTTGTGCTGCTGCTTGGCA 2182

RESULT 14

US-10-101-464A-458

; Sequence 458, Application US/10101464A

; Patent No. 6768041

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 3600
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-458

Alignment Scores:
Pred. No.: 234 Length: 3600
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-10-101-464A-458 (1-3600)

QY 354 LeuGlnValLeuAspLeuSerTyrrhen 362

DB 598 CTTGAGTTCGACCTGCTTATAAT 624

RESULT 15

US-07-792-865D-1/c

; Sequence 1, Application US/07792865D

; Patent No. 5646247

; GENERAL INFORMATION:

; APPLICANT: John W. Barnwell, Mary W. Galinski,

; APPLICANT: Samuel P. Wertheimer

; TITLE OF INVENTION: MEROZOITE ANTIGENS LOCALIZED AT

; TITLE OF INVENTION: THE APICAL END OF THE PARASITE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby P.C.

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage

; COMPUTER: IBM or IBM-compatible

; OPERATING SYSTEM: PC/MS-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/792,865D

; FILING DATE: 19911004

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US90/01849

; FILING DATE: April 3, 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson, Arthur

; REGISTRATION NUMBER: 34,354

; REFERENCE/DOCKET NUMBER: 5986/14692-US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 527-7700

; TELEFAX: (212) 753-6237

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3763 base pairs

```
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: P.vivax
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY: Lambda gt 11 native P.vivax
; LIBRARY: DNA expression library
; CLONE: 5.3
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: This sequence corresponds to
; OTHER INFORMATION: Figure 1A (sheets 1-4) in the
; OTHER INFORMATION: application, as filed.
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-792-865D-1

Alignment Scores:
Pred. No.: 245 Length: 3763
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x US-07-792-865D-1 (1-3763)

QY 730 lleLeuLeuAsnSerLeuCysPheLeu 738
Db 2576 ATTTGCTTAACCTACTATGTTTCITG 2550

Search completed: July 12, 2005, 13:02:23
Job time : 436 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:33 ; Search time 33 Seconds
(without alignments)
2051.718 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGVLLSLPVLQLAT.....PAYFTSCHLSLVAFVCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/protdata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/protdata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/protdata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/protdata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/protdata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/protdata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	4	US-09-170-496D-264
2	4596	99.9	907	4	US-09-170-496D-278
3	1217.5	25.9	644	2	US-08-866-757-2
4	1217.5	25.9	644	3	US-09-153-593-2
5	636	13.5	692	3	US-07-757-342D-6
6	636	13.5	692	4	US-09-461-657B-6
7	621.5	13.2	695	1	US-08-487-886-2
8	621.5	13.2	695	3	US-08-482-855-2
9	621.5	13.2	695	3	US-08-474-986-2
10	598	12.7	764	3	US-07-741-453A-61
11	595	12.7	764	3	US-07-757-342D-5
12	595	12.7	764	4	US-09-461-657B-5
13	595	12.7	1181	4	US-09-826-509-587
14	594	12.6	764	3	US-07-741-453A-59
15	593	12.6	764	4	US-09-826-509-403
16	591	12.6	764	3	US-07-741-453A-54
17	588.5	12.5	700	3	US-07-757-342D-3
18	588.5	12.5	700	4	US-09-461-657B-3
19	588	12.5	764	4	US-09-826-509-399
20	588	12.5	764	4	US-09-826-509-407
21	586	12.5	764	4	US-09-826-509-395
22	586	12.5	764	4	US-09-826-509-415
23	579	12.3	764	4	US-09-826-509-411
24	575	12.2	764	4	US-09-826-509-419
25	574	12.2	696	3	US-07-757-342D-4
26	574	12.2	696	4	US-09-461-657B-4
27	573	12.2	764	4	US-09-826-509-423

28	569	12.1	764	3	US-07-741-453A-60	Sequence 60, Appl
29	565	12.0	764	4	US-09-186-350A-53	Sequence 53, Appl
30	560	11.9	699	3	US-07-757-342D-2	Sequence 2, Appl
31	560	11.9	699	4	US-09-461-657B-2	Sequence 2, Appl
32	550.5	11.7	674	3	US-07-757-342D-10	Sequence 10, Appl
33	550.5	11.7	674	4	US-09-461-657B-10	Sequence 10, Appl
34	503.5	10.7	795	3	US-07-741-453A-55	Sequence 55, Appl
35	499.5	10.6	792	3	US-07-741-453A-56	Sequence 56, Appl
36	471	10.0	764	3	US-07-741-453A-29	Sequence 29, Appl
37	469.5	10.0	636	3	US-07-757-342D-7	Sequence 7, Appl
38	469.5	10.0	636	4	US-09-461-657B-7	Sequence 7, Appl
39	460	9.8	611	3	US-07-757-342D-8	Sequence 8, Appl
40	460	9.8	611	4	US-09-461-657B-8	Sequence 8, Appl
41	426.5	9.1	603	1	US-08-190-804A-50	Sequence 50, Appl
42	426.5	9.1	603	3	US-08-477-346-50	Sequence 50, Appl
43	426.5	9.1	603	4	US-08-473-089-50	Sequence 50, Appl
44	426.5	9.1	603	4	US-08-487-072A-50	Sequence 50, Appl
45	421	9.0	605	3	US-09-063-950-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-170-496D-264
; Sequence 264, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 264
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-264

Query Match	100.0%	Score 4702;	DB 4;	Length 907;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 907;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDTSLGVLLSLPVLQLATGSGSPRSVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60	
Db	1	MDTSLGVLLSLPVLQLATGSGSPRSVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60	
QY	61	PSNLVSFTSYDLSNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLUM	120	
Db	61	PSNLVSFTSYDLSNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLUM	120	
QY	121	LQNNQLRVPTAQLNLSLRDANHSIVPPSCFSGLSLRLHLLMDNALTEIPVQ	180	
Db	121	LQNNQLRVPTAQLNLSLRDANHSIVPPSCFSGLSLRLHLLMDNALTEIPVQ	180	
QY	181	AFRSLALQAMTLANKIHHIPDYAFGNLSLVLHNNRHHISLGKCFDGLHSLETLD	240	
Db	181	AFRSLALQAMTLANKIHHIPDYAFGNLSLVLHNNRHHISLGKCFDGLHSLETLD	240	
QY	241	LYNNNLDEFTAIRTLNKLGHFNHNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA	300	
Db	241	LYNNNLDEFTAIRTLNKLGHFNHNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA	300	
QY	301	FQHLPELRTLNGASQITEPDLTGATNLSLTGTGAQISLSPQVTCNLPNLOVLDLS	360	
Db	301	FQHLPELRTLNGASQITEPDLTGATNLSLTGTGAQISLSPQVTCNLPNLOVLDLS	360	
QY	361	YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDYTFQQLSLRSLNLANWKIAIHNAPST	420	

[illegible]

RESULT 2
 US-09-170-496D-278
 ; Sequence 278, Application US/09170496D
 ; Patent No. 6555339
 ; GENERAL INFORMATION:
 ; APPLICANT: Behan, Dominic P.
 ; APPLICANT: Chalmers, Derek T.
 ; APPLICANT: Liaw, Chen W.
 ; TITLE OF INVENTION: Human G Protein-
 ; ACTIVELY ACTIVATED CONSTITUTIVELY
 ; ACTIVATED HUMAN G PROTEIN-
 ; ACTIVELY ACTIVATED HUMAN G PROTEIN-
 ; ACTIVELY ACTIVATED HUMAN G PROTEIN-
 ; ACTIVELY ACTIVATED HUMAN G PROTEIN-
 ; FILE OF INVENTION: Receptors
 ; FILE REFERENCE: AREN-0040
 ; CURRENT APPLICATION NUMBER: US/09/170,496D
 ; CURRENT FILING DATE: 1998-10-13
 ; NUMBER OF SEQ ID NOS: 294
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 278
 ; LENGTH: 907
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-170-496D-278

	Query Match	99.9%	Score	4696;	DB	4;	Length	907;	
	Best Local Similarity	99.9%;	Fred.	No.	0;				
	Matches	906;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps
Qy	1	MDTSLRGVLLSLPVLLQATGCGSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSLSEL	60						
Db	1	MDTSLRGVLLSLPVLLQATGCGSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSLSEL	60						
Qy	61	PSNLSVFTSYLDLSMNNISQLLPNPLSRFLFEELRLAGNALTYTPKGAFTGLYSKVLVM	120						
Db	61	PSNLSVFTSYLDLSMNNISQLLPNPLSRFLFEELRLAGNALTYTPKGAFTGLYSKVLVM	120						

RESULT 3
US-08-866-757-2
; Sequence 2, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71.1)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA

ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2

Query Match 25.9%; Score 1217.5; DB 2; Length 644;
Best Local Similarity 46.3%; Pred. No. 6.1e-89;
Matches 250; Conservative 88; Mismatches 183; Indels 19; Gaps 5;
QY 318 ITPEDLTGTANLESILTTGAQISSLPQVCNQLPNQLVLDLSYNNLEDPSPFSCQKIQ 377
Db 2 VQOPNLGTGVHLESILTTGTAKISSIPNNLCQEQKMLRTLDLSYNNIRDLPSFNGCHALE 61
QY 378 KIDLRHNEIYEIKVDTFQQLSLRSLNLAWNKIAIHPNAPSTLPSTLKLIDSSNLLSSF 437
Db 62 EISLQRNQIYQKEGTFQGLISLRILDSRLNLIHHSRAPATLGPITNLDVSPNELTSF 121
QY 438 PITGLHGLTHLKTGNHALQSLISSENPELKVEMPYAYQCCAFGVCENAYKSNQWNK 497
Db 122 PTEGLNGLNQLKLVGNFKLKEALAAKDFVNLRSLSVPYAYQCCAFWGCDSYANL---NT 177
QY 498 GDNSSMDLHKDAGMFOQDERDLEDFLDPEEDLKALHS---VQCSPSPGPKPCEHL 554
Db 178 EDNSLQDHSVAGKGTADAANVT---TLENEE-----HSQIIHCTPSTGAFKPCBYL 228
QY 555 LDGMLIRIGVWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVNMLTGVSSAVLA 614
Db 229 LGSWMIRLTVWFIFLVALFFNLLVLTTFASCTSLPSSKLFGLISVNLPMGIYTGILT 288
QY 615 GVDATFGSFARHGAWENGVCVIGFSLIPASESSVFLTLAALERGFVSYSKAPET 674
Db 289 FLDAVSWGRFAEFGLWETGSGCKVTGFLAVFSSESAIFLLMLATVERSLAKDIMKNGK 348
QY 675 KAPFSSLVIIILLCALLATMAAVPLGSGKYGASPLCLPLPFGPEPSTGMYVALILNS 734
Db 349 SNHLKQFRVAALLAFGLATVTCGPFPHRGYSASPLCLPPTGTPSLGFTVTVLLNS 408
QY 735 LCFLLMTIATYKLYCNLDKGLDENIWDGSMVKHIALLLFTNCILNCPVAFLSFSSILNLT 794
Db 409 LAFLLMAYITYKLYCNLEKEDLSNSQSMIKHVAWLFTNCIFPCPVAFPSFAPLITAI 468
QY 795 FTSPEVIXFILLVVPVLPACLNPLIYILFNPHFKEDLSLRKQTYVYVTRSKHPSLMSINS 854
Db 469 SISPEIMKSVTLIFFPFLPACLNPLVLYVFNPKFKEDWKLKRRV---TKKSGSVSVSISS 525

RESULT 4
US-09-153-593-2

Sequence 2, Application US/09153593A
Patent No. 6174994
GENERAL INFORMATION:
APPLICANT: ELSHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DEREK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
FILE REFERENCE: GH-70055-1
CURRENT APPLICATION NUMBER: US/09/153,593A
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/866,757
EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 644
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-153-593-2

Query Match 25.9%; Score 1217.5; DB 3; Length 644;
Best Local Similarity 46.3%; Pred. No. 6.1e-89;
Matches 250; Conservative 88; Mismatches 183; Indels 19; Gaps 5;
QY 318 ITPEDLTGTANLESILTTGAQISSLPQVCNQLPNQLVLDLSYNNLEDPSPFSCQKIQ 377
Db 2 VQOPNLGTGVHLESILTTGTAKISSIPNNLCQEQKMLRTLDLSYNNIRDLPSFNGCHALE 61
QY 378 KIDLRHNEIYEIKVDTFQQLSLRSLNLAWNKIAIHPNAPSTLPSTLKLIDSSNLLSSF 437
Db 62 EISLQRNQIYQKEGTFQGLISLRILDSRLNLIHHSRAPATLGPITNLDVSPNELTSF 121
QY 438 PITGLHGLTHLKTGNHALQSLISSENPELKVEMPYAYQCCAFGVCENAYKSNQWNK 497
Db 122 PTEGLNGLNQLKLVGNFKLKEALAAKDFVNLRSLSVPYAYQCCAFWGCDSYANL---NT 177
QY 498 GDNSSMDLHKDAGMFOQDERDLEDFLDPEEDLKALHS---VQCSPSPGPKPCEHL 554
Db 178 EDNSLQDHSVAGKGTADAANVT---TLENEE-----HSQIIHCTPSTGAFKPCBYL 228
QY 555 LDGMLIRIGVWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVNMLTGVSSAVLA 614
Db 229 LGSWMIRLTVWFIFLVALFFNLLVLTTFASCTSLPSSKLFGLISVNLPMGIYTGILT 288
QY 615 GVDATFGSFARHGAWENGVCVIGFSLIPASESSVFLTLAALERGFVSYSKAPET 674
Db 289 FLDAVSWGRFAEFGLWETGSGCKVTGFLAVFSSESAIFLLMLATVERSLAKDIMKNGK 348
QY 675 KAPFSSLVIIILLCALLATMAAVPLGSGKYGASPLCLPLPFGPEPSTGMYVALILNS 734
Db 349 SNHLKQFRVAALLAFGLATVTCGPFPHRGYSASPLCLPPTGTPSLGFTVTVLLNS 408
QY 735 LCFLLMTIATYKLYCNLDKGLDENIWDGSMVKHIALLLFTNCILNCPVAFLSFSSILNLT 794
Db 409 LAFLLMAYITYKLYCNLEKEDLSNSQSMIKHVAWLFTNCIFPCPVAFPSFAPLITAI 468
QY 795 FTSPEVIXFILLVVPVLPACLNPLIYILFNPHFKEDLSLRKQTYVYVTRSKHPSLMSINS 854
Db 469 SISPEIMKSVTLIFFPFLPACLNPLVLYVFNPKFKEDWKLKRRV---TKKSGSVSVSISS 525

RESULT 5
US-07-757-342D-6
Sequence 6, Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
MINEGISHI, Takashi
NAKAMURA, Kazuto
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &


```
Db 169 -----SFESVILWLS-----KNGIEETHNCAFNG-----192
Qy 304 LPELRRLTLNGASQITEPDLTGTANLESILTLTGAQISLSPOTVCNQLPNQVLDLSYNL 363
Db 193 -----TQDELNLSD--NNN 205
Qy 364 LBDLPSPVQCQKQKIDLRHBIYEIKVDTFOQLLSRLSLNLANWKIAIHPNASTPLPS 423
Db 206 LBELEN-----DVFQAGS 219
Qy 424 LKLDLSSNLLSPPTGLHGLTHLKTGNHALQSLISSENPPELVKVIEMPYAYOCCARG 483
Db 220 PVILDISRTKVSLSLPHNGLNKLGRARSTYRKLKLPNDKFTVMEASLTYPSPHCCAPA 279
Qy 484 -----VCENA---YKISNOWNKGDN--SSMDD--LHKDQAGMFOQDERDLEDF 525
Db 280 NLKQRISELHPICNKSILRQDIDDMTQIGDQVSLIDDEPSYKSGSDMMYNEDFDLCNE 339
Qy 526 LLDPEEDLKALHSVQSPSPGPKPCHEHLLDGLIRIGVWTIAVLATCNALVTSTVFRS 585
Db 340 VVD-----VTCSPKPDAPNCPEDIMGYNILRVLIWFISILAITGNTTVLVLTTS 389
Qy 586 PLYISPIKLLIGVIAVNMVLTGVSSAVLAGVDAPTFGSPARHGAWENGVCCHVIGFLSI 645
Db 390 QYKLTVPRELMCNLAPADLCIGIYLLLIASVDIHTKSQYHNYAIDMQTGAGCDAAAGFTV 449
Qy 646 FASESVFLLTAAALRERGFYSKAFETKAPFSSLKVIILLCALLTMAAVALPGLGSK 705
Db 450 FASELSVYTLTITLERTWHITTHAQLECKVQRHAAVWVLTGTFAPAAALFPFIFGISS 509
Qy 706 YGASPLCLPLPFCPESTMGVVALILLNSLCFLMNTIATKLYCNLDKGL-ENIWDCSM 764
Db 510 YMKVSICLPMDIDSPLSQLYVMAVLLVNLVAFVWICGVYTHIVLTVRNPTIVSSSDTKI 569
Qy 765 VGHIALLLPTNCILNCPVAFSSSLINFTFISPEVIKILLVVPPLPACINPLIYLEN 824
Db 570 AKRMATLIPTDFLCNAPISFAISASLKVPLITVSKAKILLVLPFINSKANPFLYAIPT 629
Qy 825 PHFKEDLVSL-----RKOTVVTWTRSKHPSLMSINSDDVEKQSCDSTQALVT 870
Db 630 KNFRDRDFLLSKPGYENQAOIYRTETSSATNTH-----ARKSHCSSAPRV-- 677
Qy 871 FTSSITYDLPPSS 884
Db 678 -TNSYVLVPLNHSS 690
RESULT 7
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 574448
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Vui Yen
; APPLICANT: Nugent, No. 574448een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; TITLE OF INVENTION: Hormone Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Area-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 695
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: -17 to -1
; IDENTIFICATION METHOD: hydrophobic
; FEATURE:
; NAME/KEY: putative amino-terminal extracellular domain
; LOCATION: 1 to 349
; IDENTIFICATION METHOD: similarity with other
; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
; IDENTIFICATION METHOD: domains, hydrophilic
; FEATURE:
; NAME/KEY: transmembrane domain
; LOCATION: 350 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
; FEATURE:
; NAME/KEY: putative transmembrane region I
; LOCATION: 350 to 370
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region II
; LOCATION: 382 to 404
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region III
; LOCATION: 427 to 448
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region IV
; LOCATION: 469 to 491
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region V
; LOCATION: 512 to 533
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region VI
; LOCATION: 557 to 580
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region VII
```

```
; LOCATION: 592 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative carboxy-terminal intracellular
; NAME/KEY: domain
; LOCATION: 614 to 678
US-08-487-886-2

Query Match 13.2%; Score 621.5; DB 1; Length 695;
Best Local Similarity 22.4%; Pred. NO. 4.5e-41;
Matches 207; Conservative 146; Mismatches 306; Indels 265; Gaps 25;

Qy 8 VLLSLPVLQATGGSSPRSGVLLRGCTHCH---CEPDGRMLLEVDSCDGLGSELPSNL 64
Db 4 LLSLLAFSLGSG-----CHRRICHCSNRVFL---COESKVTIEPS-- 42

Qy 65 SVFTSYLDSMNISQLPNPLPLSLRFLLELRAGNALTYIPKGAFTGLYSL-KVLMQN 123
Db 43 -----DLPNAL-----ELRFVLTKLRLVQKGAFGFGDLEKIEISQN 80

Qy 124 NQLRHVPTEALQNLRSLSRLD-ANHISYVPPSCFSGLSLRLHLWLDNALTEIPVOAF 182
Db 81 DVLEIVADVFNLPKLEIRIEKANNLLYNP-----EAF 116

Qy 183 RSLSLQAMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLLETLDLN 242
Db 117 QNLNPLQVLLISNTGKHLDPV-----HKIHSLOK-----VLLDIQ 152

Qy 243 YNNLDEPPTAIRTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIQFVGRSAFQ 302
Db 153 DN-----INHTIERNFVGLS-----FESVILMLNKGIGQ 183

Qy 303 HLPRLTLTLNGASQITEFPDLTGANLESULTGCAQISSLPQVCNQLPNQLVDLSYN 362
Db 184 --EIHNCAFNG-----TQDELNLSD--NN 204

Qy 363 LLEDLPFSVCQKQKIDLRHNEIYEIKVDYFQQLLSRLSLNLANWKIATHPNAPSTLP 422
Db 205 NLEELPN-----DVFHGAS 218

Qy 423 SLIKDLSSNLLSSPITGLHGLTHLKLGTGNHALQSLISSENFPELKVIEPMPYAYQCCAF 482
Db 219 GPVILDISRTIRHSLPSVGLNKLKLRASYNLKPLTLEKLVALMEASLTYPHCCAF 278

Qy 483 G-----VCENAYKISNQ-----WNGDNSSMDLDLHKDAGMFOQDERDLDP 525
Db 279 ANWRQIISLHPICNKS--ILRQEVYDYMVTQTRQSRSLAEDN-----ESSYKRGFDMT 329

Qy 526 LLDPEEDL-KALHSVQCSPPKPCHEHLLDGLWLRIGVMTIATVLAALTCNALVTSTVFR 584
Db 330 YTEFDYDLCNEVVDYTPKPDAPNCPEDIMGYNILRLWIFISILAITGNIILVLVILT 389

Qy 585 SPLYSPIKLIGVIAAVNMLTGVSVALAGVDAFTFGSGFARHGAWENGVCCHVIGFLS 644
Db 390 SOYKLTVPFLMCNLAFADLCIGIYVILLIASVDIHTKQYHNYADWQTGACDAAGFT 449

Qy 645 IFASSESVFLFLTAALERGFYSVKYSAKPTETKAPFSLSKVIILLCALLALTWAAPVLLGGS 704
Db 450 VFASLSVVVTLTAITLERWHITTHAMQLDCKVQLRHAASVMVMGWFIFAPAAALPFIIGIS 509

Qy 705 KYGASPLCLPFPFGPSMGYNVAILLNSLCFLAMTAYTKLYCNLDKGLD-ENIMDCS 763
Db 510 SYMKVSIICLPMIDISPLSOLYMSLVNLVNLAFVVICGYTHIYLTVRNPVIVSSSDTR 569

Qy 764 MVKHALLLFTICNLCPVAFPSFLINLTIPISPEVIKFIILVVVPLPACINPLLYLILF 823
Db 570 TAKRWAMLIFTDFLCWAFISFAISLKVPLITVSKAKILLVFLHPINSCANPFLYAI 629

Qy 824 NPHFKEDLVSLRKQTYVWTRSKHPSLMSINSDVEKQSCDSTQALV--TFTSSSTYDLP 881
Db 630 TKNFRDRFEL-----LSKCGCYEQAOIYRTETSSTVHNTHP 667
```

```
Qy 882 PSSVSPAYPVPVPTESCHLSSVAFVP 905
Db 668 RGHCSAPRVTVNG--STYLVP 688

RESULT 8
US-08-482-855-2
; Sequence 2, Application US/08482855
; Patent No. 6121016
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Vui Yen
; APPLICANT: Nugent, No. 6121016een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; TITLE OF INVENTION: Hormone Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,855
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 695
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: -17 to -1
; IDENTIFICATION METHOD: hydrophobic
; FEATURE:
; NAME/KEY: putative amino-terminal extracellular domain
; LOCATION: 1 to 349
; IDENTIFICATION METHOD: similarity with other
; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
; IDENTIFICATION METHOD: domains, hydrophilic
; FEATURE:
; NAME/KEY: transmembrane domain
; LOCATION: 350 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
; FEATURE:
; NAME/KEY: putative transmembrane region I
; LOCATION: 350 to 370
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
```

:	NAME/KEY:	putative transmembrane region II
:	LOCATION:	382 to 404
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative transmembrane region III
:	LOCATION:	427 to 448
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative transmembrane region IV
:	LOCATION:	469 to 491
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative transmembrane region V
:	LOCATION:	512 to 533
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative transmembrane region VI
:	LOCATION:	557 to 580
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative transmembrane region VII
:	LOCATION:	592 to 613
:	IDENTIFICATION METHOD:	similarity to other G
:	IDENTIFICATION METHOD:	protein-coupled receptor transmembrane regions,
:	IDENTIFICATION METHOD:	hydrophobic, about 20-23 amino acids in length
:	FEATURE:	
:	NAME/KEY:	putative carboxy-terminal intracellular
:	NAME/KEY:	domain
:	LOCATION:	614 to 678
:	US-08-482-855-2	
 Query Match 13.2%; Score 621.5; DB 3; Length 695;		
Best Local Similarity 22.4%; Pred. No. 4.5e-41;		
Matches 207; Conservative 146; Mismatches 306; Indels 265; Gaps 25		
Qy	8	VLLSPVLQLATGGSSPSRGVLLRGCPHCH---CEPDGRMLLRVDSCDLGLSELPSNL 64
Dd	4	LAVSLIAFLSLGSG-----CHRRICHCSNRVFL---CQESKVTETPS-- 42
Qy	65	SVFTSVLDLSMNNISQLPNPLPSLRPLEELRAGNALTYIPKGAFITGLYSL-KVLMLOJN 123
Dd	43	-----DLPRNAI : : : : : : : :
Qy	124	NOLRHVPTEALQNRLSLOSRLID-ANHISYYPPSCFSGLHSRLHLWDDNALTEIVPQAF 182
Dd	81	DVLEIEADVPSNPDKHEIRIEKANLLYNP-----EAF 116
Qy	183	RSLSALOAMTLANKIHIIPOYAFGNLSLVVLHLHNRIHSLGKKCFDGLHSLETLDLN 242
Dd	117	QNLPNLQYLIIISNTGIKLHPDV-----HKIHSLQK-----VLLDIQ 152
Qy	243	YNNLDEFPPTAIRTLSNLKELGPHSNNIRSIPKAPGVGNPSLIITHFYNDPIQFVGSRSAQ 302
Dd	153	DN-----INHIITEENSFVGLS : : : : : : : : : :
Qy	303	HLPFLRTLTINGASQITEPFDLTGTANLESRTLTAQISSLIPQTVCNQLPNLQVLDLSYN 362
Dd	184	---EIHNCAFG-----TQDELNLSD--NN 204
Qy	363	LLEDLPFSFCVKQLOKIDLRHNEIYEIKVDTFQQLLSRSNLNLAWNKIALIHPNAESTJP 422
Dd	205	NUELEFN-----DVFHGAS 218

	Qy	720 PSTWGYVVALILLNSLCFLMTTIATYKUYCNL-----DKGDLNWTDCSMGVHIAALLFT 774
	Dd	577 PLALAYIVFVLITNAIEVAEIVCCCVKIYIIVRNPOYNPGDK----DTKIAKRMAVLIFT 632
	Qy	775 NCILNCPAVELSFSSLINLTETISPEVIKFILLVVPLPACLNPLAYLIFNFHFKEIDVSL 834
	Dd	633 DFICMAPISFYALSAILNKPLITVSNKSILIVLFYPNLSCANPFYAIFTEKFORDVFIL 692
	Qy	835 -----RKQTYVVTNRKHPSILMSINDSDVEKQSCDSTQAALTFTTSSSIYYDLPSPSSVPS 887
	Dd	693 LSKFGICKRQAQRGORVPKP--NSTDIQVOK-----VTHMRQGLHNH-----E 736
	Qy	888 PAYPVTESCHLS 899
	Dd	737 DYEELIEKSHLT 748

RESULT 12
 US-09-461-657B-5
 ; Sequence 5, Application US/09461657B
 ; Patent No. 6635445
 ; GENERAL INFORMATION:
 ; APPLICANT: IGARASHI, MASAO
 ; APPLICANT: MINEGISHI, TAKASHI
 ; APPLICANT: NAKAMURA, KAZUTO
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN LUTEINIZING
 ; HORMONE-HUMAN CHORIONIC GONADOTROPIN RECEPTOR PROTEIN
 ; TITLE OF INVENTION: AND TRANSFORMANTS THEREOF
 ; FILE REFERENCE: 70281/41,226C
 ; CURRENT APPLICATION NUMBER: US/09/461,657B
 ; CURRENT FILING DATE: 2002-04-16
 ; PRIOR APPLICATION NUMBER: 07/757,342
 ; PRIOR FILING DATE: 1991-09-10
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 764
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-461-657B-5

Query Match	12.7%;	Score 595;	DB 4;	Length 764;
Best Local Similarity	21.9%;	Pred. No. 6.9e-39;		
Matches 213;	Conservative 148;	Mismatches 301;	Indels 310;	Gaps 31;
Qy	9	LLSPVLLQL-----ATGGSSPRSGVLLRGCPPTHCHCEPDMGLRLRVDCSLGSLSELP	63	
Db	6	LLQVLLLLDLPRDLGGWCSSP-----PCEHQEED-----FRVTKDI-----	44	
Qy	64	LSVFYSYDLSMMNNISQLPNPLPSLRPLEELRLAGNALTYIPKGAFTGLYSKLVMLQN	123	
Db	45	-----QRIPLSP-----	52	
Qy	124	NQLRHVTEALQNLRSLOSLDANHHISYVPPSCFSGLSHLRLWLDNALTEIPVQAFR	183	
Db	53	-----STQTLKIETHURTIESHAPSNLPNISRIYVS-----IDVTLOOLE	93	
Qy	184	SLSLAQMTALNKKIHHIPDYAFGNLSLVVLHLHNNRIHSLGKKCFDGLHSLETLDINY	243	
Db	94	S-----HSFYNLKSVTHIEIRNTR-----	114	
Qy	244	NNLDEFFTAIRTLNKLKGHSHNNIRSIPEKAFVGNPSLTIHYPNDPIQFVGRSAPQH	303	
Db	115	TYID--PDALKELPLKFLGINTGLKMPDITKV-----YSTDIEFI-----	155	
Qy	304	LPELRTTLNGASQITEPPDITGTANLESRTLGAQISSLP-----QTCNQLPNLVLIDL	359	
Db	156	-----LEITDIPYMT-----SIPWNAFOGLCNETITLKLKLYNN	187	
Qy	360	SYNNLEDLPFSVCQKQKIDLRNNEIYEIKVDTFQQLLSLRSLNLANWKIAIHPNAPS	419	
Db	188	GFTSVQGY-AFN-GTKLDAYLVNKKY-----LTVIYKDAFG	222	

Qy	420	TL---	PSLIKLDLSNLLSFPITGLHGLTHLTKTGNHALQSLISSENFPFLKVIEMPYA	476				
Db	223	GVYSGPFL--LDVSGT	SVTALPSKGLSHLKELIARNTWTTLKGLPLSLFHLTRADLSY	280				
Qy	477	YQCAFG-	--VCENAYKLSNOWNKGDNSMDLHK-----KDA	511				
Db	281	SHCCAFKNQKKIRIGILES	MCNTESSQSLRQKSVNALNSPLHREYENLGDSIVGYKEK	340				
Qy	512	GMFQAQDE-----	RDLEDFLDF-----EEDLKALHS-----VQCS	543				
Db	341	SKQDTHNNAHYVFFEEQ	DEIIGFQGLKNPQEEITLQAFD	SHVDVYTCGDS	EDMVC	TP	400	
Qy	544	SPGFPKPEHLLDGLWIRIGW	TIANVALTCNALVTSPRSP	LIYISPIKLLIGWIAAVN	603			
Db	401	KSDFNPCEDIMGYKFLRI	YVWFSSLLALGNVFVLLIL	LTSHYKLVNPRELMCNLAF	460			
Qy	604	MLTCVSSAVLAGYDAFTFGS	FARHGAWENGVC	HVIGFELSIFASESSVFLTL	LAALBERG	663		
Db	461	FCMGMYLLLLASVDLYTHSE	YNNHAIQWQPGCNTAG	FTFVFASELSVYTL	WTILBERW	520		
Qy	664	FSVKYSAKFTKAPFSSLK	VII---LLCALLAL	TMAAVPLVGGSKYGASPLCLPLP	PEGE	719		
Db	521	YATFAMBLDRKMLRHACA	IMVGGVCCFLAL---	LPLVGISSYAKVSGICLPM	DTET	576		
Qy	720	PSWGYVVALILNSLCFLM	WTIATYKLYCNL-----	DKGDLNIDCWSWKHALLFT	774			
Db	577	PLALAYTVFVLTINIAV	FIVCCYKXIYITVRNPQ	INPGDK-----DTKIAKRAV	LAFLFT	632		
Qy	775	NCILNCVPAFLSPSSLLN	LFI	SEVTKFILLVVPPLACP	LNPLLYTLFNP	HFKED	VS	834
Db	633	DFICMAPISYALSAILN	KPLITVSN	SKILLVLFYPLNSCAN	PLFYALFT	KEFOR	DVIL	692
Qy	835	-----RKQTYVW	TRSKHPSLMSIN	DDVEKQSCDSTQ	ALVFTTSSSIT	YDLPLPSS	VPS	887
Db	693	LSKFGICKRQAQYRG	QRPVK---NSTD	IQVQK-----VTHMRQGLHN	-----E	736		
Qy	888	PAYPVTESCHLS	899					
Db	737	DVYELIEKSHLT	748					

```

RESULT 13
US-09-826-509-587
; Sequence 587, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Liin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 587
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-587

```

```

Query Match      12.7%; Score 595; DB 4; Length 1181;
Best local Similarity 22.0%; Pred. No. 1.3e-38;
Matches 218; Conservative 145; Mismatches 303; Indels 324; Gaps 32;

Qy 9 LLSPVLVLLQ-----ATGSSPSRSGVLLRCPTHCHCEPDGRMLLRVDCSDGLGSELSPN 63
Db 6 MGLVLLDIPDLGGMGCSPP-----PCEHOED-----FRVTKDI-----44

```

```
QY 64 LSVFTSYLDLSMNNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLMLQN 123
Db 45 -----QRIPSLPP----- 52
QY 124 NQLRHVPTEALQNLRSQSLRDANHSIVYPPSCFGLSHSLRHLWLDNNALEIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIPSHAFSNLPNISRIYV-----IDVTLQOLE 93
QY 184 SLSALQAMTALANKIHHIPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLDLNY 243
Db 94 S-----HSFYNLKSVTHIEIRNTR-----NL 114
QY 244 NNLDEFPFAIRTLNMLKELGFHNNIRSIPEKAFVGNPSLTIHFDNPIQVGRSAFQH 303
Db 115 TYID--PDALKELPLKSLAFNTGLKMPFDLTKV-----YSTDIFFI----- 155
QY 304 LPELRTLINGASQITTEFPDLTGTANLESLLTGTQAQISSLP-----QTVCNQLPNLOVLDL 359
Db 156 -----LEITDNPYMT-----SIPVNAFQGLCNETLTKLYNN 187
QY 360 SYNLEDLPSFVCOVKQKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAIHPNAPS 419
Db 188 GFTSVQGYDFFGT--KLDAVYLNKKY-----LTVIDKDAFG 222
QY 420 TL---PSLIKLDLSNLLSPPIGLHGLTHLKLGTGNHALQSLSISSENPPELKVEMPIYA 476
Db 223 GYSGPSL--LDVSTSVTALPSKGLHKLKELIARNSTLTKKIALSLSLFLHLTRADLSYP 280
QY 477 YQCCAF-----GVCE-----NAYKISN-OMNKGDNSSMDLHK-----KDA 511
Db 281 SHCCAFKQKIRGILESLMNCNESSIETLRQRKSYNALNSPLHQBEBNGLGDSIVGYKEK 340
QY 512 GMFOAQDE-----RLEDPLDF-----BEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFREGEDEIIGFQELKNPQETLOAFDQSHDYDTICGSEDMVCTP 400
QY 544 SPGPFPKCEHLDOGWLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIAAVN 603
Db 401 KSDEFNPECEDIMGYKFLRIWVFWFVSLALLGNVFLVLLITLTSYKLVNVPFLMCLNAPAD 460
QY 604 MLTGSSAVLAGVADFTGSPARHCAWENGVCVHGFSLFASSESVFLLTLAALBERG 663
Db 461 FCMGYLLLIASVDLITSEYNNHAIWDQTPGCGNTAGFTVFASELSVYTLTVITLERW 520
QY 664 PSVKYSAKFETKAPFSSLVKVII-----LLCALALATMAAIVPILGGSKYGASPLCLPLPFG 719
Db 521 YAITFAMALDRKIRLRHACAIMVGWVCCFLAL-----LPLVGISSYAKVSICLPMDTET 576
QY 720 PSTMGYVVALILNLSCLFMMTIATYKLYCNL-----DKGDLNENIWDSCMVKHIALLLFTN 775
Db 577 PLALAYIVFVLTLNIVAFVIVCCVVKIYITVRNPHNPGDK-----DTKIAKEMAVLIIFTD 632
QY 776 CILNCPVAPLFSSSLINITFTSPVVKIFILLVVVPLPACPLPLLYLFPNHPKEDLVSL- 834
Db 633 FTCMAPISEYAVSAILNKPLITVSNKILLVLFYPINSCANPFLYAIPTKAFQORDVIFLL 692
QY 835 -----RQOTVYWTSEKPSLMSI-----NSDDV-----EKQS 861
Db 693 SKFGICKRQAOAIRGORVPPKNSTDIQVKVTHDMRQGLHNMEDVYELIENSHLTPPKQ 752
QY 862 CDSTQALVTFTSSSITYD-----LPPSSVPS 887
Db 753 QISEEYMQTVLAISAEPHHTGLVDPSSVPS 782
```

RESULT 14

US-07-741-453A-59
; Sequence 59, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC

APPLICANT: DUMONT, JACQUES
APPLICANT: VASSART, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-741-453A-59

Query Match 12.6%; Score 594; DB 3; Length 764;
Best Local Similarity 21.9%; Pred. No. 8.3e-39;
Matches 213; Conservative 149; Mismatches 299; Indels 312; Gaps 31;

```
QY 9 LLSLPVLQL-----ATGSSPSRSGVLLRGCTHCHCEPDGRMLLRVDCSLGLSELPSN 63
Db 6 LQLVLLLDLPRDLGGMGSSP-----PCEHQEED-----PRVTCKDI----- 44
QY 64 LSVFTSYLDLSMNNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKVLMLQN 123
Db 45 -----QRIPSLPP----- 52
QY 124 NQLRHVPTEALQNLRSQSLRDANHSIVYPPSCFGLSHSLRHLWLDNNALEIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIPSHAFSNLPNISRIYV----- 83
QY 184 SLSALQAMTALANKIHHIPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLDLNY 243
Db 84 -----SIDTLQOLE-----SHSFYNLKSVTHIEIRNTR-----NL 114
QY 244 NNLDEFPFAIRTLNMLKELGFHNNIRSIPEKAFVGNPSLTIHFDNPIQVGRSAFQH 303
Db 115 TYID--PDALKELPLKSLAFNTGLKMPFDLTKV-----YSTDIFFI----- 155
QY 304 LPELRTLINGASQITTEFPDLTGTANLESLLTGTQAQISSLP-----QTVCNQLPNLOVLDL 359
Db 156 -----LEITDNPYMT-----SIPVNAFQGLCNETLTKLYNN 187
QY 360 SYNLEDLPSFVCOVKQKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAIHPNAPS 419
Db 188 GFTSVQGY-APN-GTKLDAVYLNKKY-----LTVIDKDAFG 222
QY 420 TL---PSLIKLDLSNLLSPPIGLHGLTHLKLGTGNHALQSLSISSENPPELKVEMPIYA 476
Db 223 GYSGPSL--LDVSTSVTALPSKGLHKLKELIARNSTLTKKIALSLSLFLHLTRADLSYP 280
```

```
Qy 477 YQCCAFG-----VCENAYKISQWKNKGDNSSMDLHK-----KDA 511
Db 281 SHCAFKXQKKIRGILBSLNCNESSMQSLRQKSVNALNSPLHQEYENLGDSIVGYKEK 340
Qy 512 GMFOAQDE-----RDLEDFLDF-----EEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFEEQDEIIIGFQELKNPQETLQAFDSDHYDYTI CGSDEMVCPT 400
Qy 544 SPGFKPCHEHLLDGWLRIGVWTIAVLTALCNALVTSTVFRSPLYISPIKLLIGVIAAVN 603
Db 401 KSDEFNCPEDIMGYKFLRIVVWFVSLALLGNVFLVLLTSHYKLANVPFLMCLNAPAD 460
Qy 604 MLTGSSAVLAGVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLTLAALERG 663
Db 461 FCMGYMLLLIASVDLYTHSEYHNHAIWQTPGCGNTAGFTVFASELSVYTLVTILRW 520
Qy 664 FSVKYSKAFETKAPFSSLKVII-----LLCALLALTMAAVPLLGSKYKYGASPLCLP 719
Db 521 YAITFAMRLDRKIRLRHACAIMVGGWCCFLAL-----LPLVGISSYAKVSICLPMDTET 576
Qy 720 PSTWGYWVALILLNSCLFLMMTIATYKLYCNL-----DKGLENINWDCSWKHIALLLFT 774
Db 577 PLALAYIVFVLTINIAVAVIVCCCVKIYITVRNPQYNPGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFLSFSSLINLTISPVEIKFILLVVPPLACLNPLLYILFNPHEKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKPLITVNSKILLVLYPLNSCANPFLYAITKAFORDVIL 692
Qy 835 -----RKQTYVWTRSKHPSLMSINSDVVEKQSCDSTQALVTFTSSSITYDLPPS-SVP 886
Db 693 LSKFGICKRQAQAYRGQVPEK---NSTDIQVQ-----KVTHDMRQGLHNM 735
Qy 887 SPAYPVTESCHLS 899
Db 736 EDVYELIEKSHLT 748

RESULT 15
US-09-826-509-403
; Sequence 403, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 403
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-403
```

Query Match 12.6%; Score 593; DB 4; Length 764;
Best Local Similarity 21.9%; Pred.No.1e-38;
Matches 213; Conservative 148; Mismatches 301; Indels 310; Gaps 31;

```
Qy 9 LLSLPLVLQL-----ATGSSPSRSGVLLRGCPHTCHCEPDGRMLLRVDCSLGLSELPSN 63
Db 6 LLQLVLLLDLPRDLGGWCCSP-----PCEHQEED-----FRVICKDI-----44
Qy 64 LSVFTSYLDLSMNNTISQLPNPLPSRFLBELRLAGNALTYIPKGAFTGLYSKLKVLMLQN 123
```

Search completed: July 12, 2005, 07:49:29
Job time : 45 secs

```
Db 45 -----ORIPSLP-----52
Qy 124 NOLHVPTEALQNLRSLSQSLRLDANHLISYVPPSCFGLSHSLRHLWLDNALTEIPVQAFR 183
Db 53 -----STQTKLIETHRTIPSHAFSNLPMISRIYVS-----IDVTLQOLE 93
Qy 184 SLSALQAMTLALNKHIIHPDYAFGNLSLVLHLHNNRIHSLGKCKCFDGLSHSLTDLNY 243
Db 94 S-----HSFYNLSKVTHIERNTR-----NL 114
Qy 244 NNLDFFTARTLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPTQFVORSAPQH 303
Db 115 TYID--PDALKELPLLLKFLAGTNGKMPDLTKV-----YSTDIPFI-----155
Qy 304 LPELRTITLNGASQITEFPDITGTANLESLLTGAQISSLP-----QTVCNQLPNLOVL 359
Db 156 -----LEITDNPMYMT-----SIPVNAFQGLCNETLTKLYNN 187
Qy 360 SYNLEDLPSFVQCQKQIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIATIHNAFS 419
Db 188 GFTSVQGY-AFN-GTKLDVAVLNKXY-----LTVIDKDAFG 222
Qy 420 TL-----PSLIKLDLSSNLLSPFTTGLHGLTHLKTGNHALQSLISSNFPBELKVIENPYA 476
Db 223 GVYSGPSL--LDVSQTSVTALPSKLEHLKELIARNTWTLLKPLSLSLFLHTRADLSYP 280
Qy 477 YQCCAFG-----VCENAYKISQWKNKGDNSSMDLHK-----KDA 511
Db 281 SHCAFKXQKKIRGILBSLNCNESSMQSLRQKSVNALNSPLHQEYENLGDSIVGYKEK 340
Qy 512 GMFOAQDE-----RDLEDFLDF-----EEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFEEQDEIIIGFQELKNPQETLQAFDSDHYDYTI CGSDEMVCPT 400
Qy 544 SPGFKPCHEHLLDGWLRIGVWTIAVLTALCNALVTSTVFRSPLYISPIKLLIGVIAAVN 603
Db 401 KSDEFNCPEDIMGYKFLRIVVWFVSLALLGNVFLVLLTSHYKLANVPFLMCLNAPAD 460
Qy 604 MLTGSSAVLAGVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLTLAALERG 663
Db 461 FCMGYMLLLIASVDLYTHSEYHNHAIWQTPGCGNTAGFTVFASELSVYTLVTILRW 520
Qy 664 FSVKYSKAFETKAPFSSLKVII-----LLCALLALTMAAVPLLGSKYKYGASPLCLP 719
Db 521 YAITFAMRLDRKIRLRHACAIMVGGWCCFLAL-----LPLVGISSYAKVSICLPMDTET 576
Qy 720 PSTWGYWVALILLNSCLFLMMTIATYKLYCNL-----DKGLENINWDCSWKHIALLLFT 774
Db 577 PLALAYIVFVLTINIAVAVIVCCCVKIYITVRNPQYNPGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFLSFSSLINLTISPVEIKFILLVVPPLACLNPLLYILFNPHEKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKPLITVNSKILLVLYPLNSCANPFLYAITKAFORDVIL 692
Qy 835 -----RKQTYVWTRSKHPSLMSINSDVVEKQSCDSTQALVTFTSSSITYDLPPSVP 887
Db 693 LSKFGICKRQAQAYRGQVPEK---NSTDIQVQ-----VTHMRQGLHNM-----E 736
Qy 888 PAYPVTESCHLS 899
Db 737 EDVYELIEKSHLT 748
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 12, 2005, 21:29:25 ; Search time 197.5 Seconds
(without alignments)
10668.712 Million cell updates/sec

Title: US-10-751-736-21
Perfect score: 907
Sequence: 1 atggacacctccggctcgg.....cattgtcccatgtctctaa 2724

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Word size: 1
Total number of hits satisfying chosen parameters: 3967864

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlh
-Q=/cgn2.1/USPTO.spool/US10751736/runat.12072005.084327.23907/app.query.fasta_1.2887
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CGN 1.1 279 @runat.12072005.084327.23907 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04.*
1: Genesecp1980s.*
2: Genesecp1990s.*
3: Genesecp2000s.*
4: Genesecp2001s.*
5: Genesecp2002s.*
6: Genesecp2003as.*
7: Genesecp2003bs.*
8: Genesecp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	2	Aaw93889 Human HG3
2	907	100.0	907	3	Aay90682 Human G p
3	907	100.0	907	6	Abp81968 Human G p
4	907	100.0	907	6	Abp06467 Human G p
5	907	100.0	907	7	Adc22783 Human G p
6	907	100.0	907	7	Adc59150 Human G p
7	907	100.0	907	7	Adc59153 Human Pro
8	907	100.0	907	7	Adg42628 Human G p
9	907	100.0	907	7	Adh14256 Human HG3
10	907	100.0	907	7	Adn40013 Cancer/an

11	907	100.0	907	7	ADN39531	Adn39531	Cancer/an
12	907	100.0	907	7	ADN39628	Adn39628	Cancer/an
13	907	100.0	907	8	ADO29408	Ado29408	Human GPC
14	907	100.0	907	8	ADQ80369	Adq80369	G protein
15	907	100.0	907	8	ADR67868	Adr67868	Human HG3
16	806	88.9	907	3	Aay90687	Aay90687	Human mut
17	806	88.9	907	3	ADC22797	Adc22797	Human G p
18	806	88.9	907	7	ADH14270	Adh14270	Mutated h
19	806	88.9	1145	7	ADF70480	Adf70480	Orphan re
20	716	78.9	907	7	ADG42629	Adg42629	Human G p
21	689	76.0	692	2	AAW93890	Aaw93890	Human HG3
22	621	68.5	883	7	ADB80464	Adb80464	Ovarian c
23	621	68.5	883	7	ADN40012	Adn40012	Cancer/an
24	621	68.5	883	7	ADN39166	Adn39166	Cancer/an
25	592	65.3	693	2	AAy42169	Aay42169	Human LGR
26	202	22.3	202	2	AAy53573	Aay53573	Human gon
27	197	21.7	282	5	ABP42000	Abp42000	Human ova
28	51	5.6	907	8	ADG75449	Adg75449	Mouse ort
29	51	5.6	907	8	ADO29409	Ado29409	Mouse GPC
30	43	4.7	43	2	AAW93892	Aaw93892	Human HG3
31	43	4.7	43	2	AAW93894	Aaw93894	Human HG3
32	43	4.7	43	2	AAW93893	Aaw93893	Human HG3
33	43	4.7	43	2	AAW93895	Aaw93895	Human HG3
34	43	4.7	43	2	AAW93896	Aaw93896	Human HG3
35	43	4.7	43	2	AAW93891	Aaw93891	Human HG3
36	20	2.2	20	6	ABP82972	Abp82972	G protein
37	20	2.2	20	6	ABP82976	Abp82976	G protein
38	18	2.0	91	2	AAy53577	Aay53577	Human gon
39	18	2.0	244	8	ADO66490	Ado66490	Novel hum
40	18	2.0	266	3	AA841583	Aab41583	Human ORF
41	18	2.0	736	5	ADN02246	Adn02246	Human LGR
42	18	2.0	736	5	ADK19411	Adk19411	Human par
43	18	2.0	736	5	ADH50810	Adh50810	Human G-p
44	18	2.0	794	5	ADH50810	Adh50810	Human G-p
45	18	2.0	823	5	ADH50809	Adh50809	Human G-p

ALIGNMENTS

RESULT 1

AAW93889
ID AAW93889 standard; protein; 907 AA.

AC AAW93889;

DT 25-JUN-1999 (first entry)

DE Human HG38 protein.

DE Human HG38 protein.

KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
KW endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator.

OS Homo sapiens.

XX WO9915660-A1.

XX 01-APR-1999.

PF 24-SEP-1998; 98WO-US019979.

PR 24-SEP-1997; 97US-0059863P.

XX (MERI) MERCK & CO INC.

PI Liu Q, Bailey WJ, McDonald TP;

XX WPI; 1999-254711/21.

DR N-PSDB; AAX23980.

PT Human G-protein coupled glycoprotein hormone receptor HG38.

PS Claim 1a; Fig 2; 74pp; English.

Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
QY 241 CTGCTCCGGAATCCCTCCAGCTCCCGCTTCCTCGGAGGATTAAGTCTTCGCGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
QY 301 GCTCTGACATACATTCACAGGAGCATTCAGTGCCTTTACAGTCTTAAAGTCTTTATG 360
Db 101 AlaLeuThrTyrIleProLyseGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
QY 361 CTGCAGAAATATACGCTAAGACACCTGCCACAGAAAGCTCTGCAGAAATTTGGAGACCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
QY 421 CAATCCCTGCGCTCGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
QY 481 CTGCATTCCTGAGGCACCTGTGGCTGGATGACAAATGGTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
QY 541 GCTTTTAGAAGTTTATCGCATGTCGAAGCCATGACCTTGGCCCTCGAACAAATAATACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
QY 601 ATACACAGACTATGCTTTGGAAACCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
QY 721 TTAATATCAATAATCTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
QY 781 GAACTAGGATTTTCATAGCAACAATATCAGGTGATACCTCGAAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
QY 841 CTTCTCTTTATACATATTCATGACAAATCCATCCCATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTTACCTGAATAGAAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTCTGATTTAACTGGAACCTGCAACCTGGAGACTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
QY 1021 TCATCTCTCTCAAAACCGCTCGCAATCAGTTACCTAATCTCCAAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAAGATTTACCCAGTGTTCAGTCTGCCAAAGCTTCAGAAAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGACCTATCGTCCAACTCTCTGCTCTTTCTCTATAACT 1320

Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTTACAGTGCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAATCCCTATAGAATTTCTTAATCAATGAATPAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGAGCAGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAGATCAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCCCTGCTTACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTACACCTTCCCAGGCCCCCTTCAAAACCTGTGAAACCTGCTGTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATGGAGTGTGGAACCATAGCAGTCTGCGACTTACTTGTATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTTAAATGGGGTCAATCCCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCGCTGCTGGTGGTGGATGCGTTCATC 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGGTGGTGGTGCATGTCAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTGTCCATTTTGTCTCAGATCATCTGTTTCTGCTTCTTACTCTGCGACGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAAATCATTTTGTCTGTGCCCTGTGCGCTTGACCATGGCCGCGAGTTCCTCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAAGATGGCGCTCCCTCTCTGCGCTGCTTTCCTTTTGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGGTACATGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTCGAGAAATTTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTAGGTAAACACATTCCTGCTTCTTCCAACTGATCCTTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGGCTCTTCTGCTCTCTCTTAAATAAACTTACATTTATCAGTCTCCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800

|||||
181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
|||||
601 ATACGAGACTATGCTTTGGAAACCTCTCCAGCTGGTAGTCTACATCTCCATAACAAT 660
|||||
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuHisLeuHisAsnAsn 220
|||||
661 AGAATCCACTCCCTGGGAAGAAAGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
|||||
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGlnThrLeuAsp 240
|||||
721 TTAATAATCAATAAACCCTTGATCAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA 780
|||||
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
|||||
781 GAACATAGATTTATAGCAACAATATCAGGTGATCTGATCTGAGAAAGCAATTTGTAGGCAAC 840
|||||
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
|||||
841 CCTTCTCTATTACAAATCTTATGACATCCATCCATTTGTTGGGAGATCTGCT 900
|||||
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
|||||
901 TTTCAACATTTACCTGAACATAAGAACACTGACTCTCAATGGTGCCTCACAATAAAGTCAA 960
|||||
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
|||||
961 TTTCTGTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
|||||
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
|||||
1021 TCATCTCTCTCAACCTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGCT 1080
|||||
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
|||||
1081 TACAACCTATTAGAAGATTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC 1140
|||||
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
|||||
1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTGTTAGCTTC 1200
|||||
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
|||||
1201 CGATCGCTGAATTTGGCTGGACAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
|||||
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
|||||
1261 TTGGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGTCGTTTTCCTATAACT 1320
|||||
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
|||||
1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
|||||
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
|||||
1381 TCATCTGAAACTTTCCAGACTCAGGTTATAGAAATGCTTATGCTTACCAGTCTGCT 1440
|||||
461 SerSerGluAsnPheProGluLeuLysValIleGluMecProTyrAlaTyrGlnCysCys 480
|||||
1441 GCATTTGGAGTGTGTGAGAAATGCTATAAGATTCTTAATCAATGGAAATAAAGTGACAAC 1500
|||||
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
|||||
1501 AGCAGTATGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGACGT 1560
|||||
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
|||||
1561 GACCTTGAGATTTCTGCTGCTGCTTTGAGGAGACCTGAAAGCCCTTCATTACAGTGCAG 1620
|||||
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
|||||
1621 TGTTCACCTTCCCAGGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC 1680
|||||

541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1681 AGAATTGGAGTGGACCATAGCAGTCTCTGCACTTACTTCTAATGCTTTGGTGACTTCA 1740
|||||
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
|||||
1741 ACAGTTTTTCAGATCCCTCTCTGTACATTTCCCAATTAACCTGTTAATTTGGGGTCAATCGCA 1800
|||||
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
|||||
1801 GCAGTGAACATGCTCACGGAGATCTCCAGTCCCTGCTGGCTGGTGTGCTGAGTCCGTTCACT 1860
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
|||||
1861 TTTGCGACGTTTGCACGACATGGTGGCTGGGAGAAATGGGTTGGTGGCATGCTCAAT 1920
|||||
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
|||||
1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
|||||
1981 GAGCTGGGTTCTCTGTGAAATATTTCCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
|||||
2041 CTGAAAGTAATCATTTTGTCTCTGTCCTCTGTCCTTGGCCTTGGCATGCGCGCAGTCCCTCG 2100
|||||
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
|||||
2101 CTGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGCCC 2160
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
|||||
2161 AGCACCATGGCTACATGCTGCTCTCATCTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
|||||
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGAACAGGAGAGCTCGAGAAATATTTGG 2280
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
|||||
2281 GACTCTCTATGGTAAAAACACATTCCTGCTTCTTCCAACTGCATGCAATCTAAACTGC 2340
761 AspCysSerMetValLysHisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCys 780
|||||
2341 CCTGTGGCTTTCTGCTCTCTCTCTTAAATTAACCTTACATTTATCAGTCTCTGAAGTA 2400
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
|||||
2401 ATTAAGTTTATCTTCTGCTGCTAGTCCCACTTCTCTGATGCTCAATCCCTTCTCTAC 2460
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
|||||
2461 ATCTTGTTCATPCTCCTTTAAGAGGATCTGGTGGCTGAGAAAGCAAACTTACGTC 2520
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
|||||
2521 TGGACAGATCAAAACACCCAGCTGATGTCATTAATTACTCTGATGATCTCGAAAACAG 2580
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
|||||
2581 TCCTGTGACTCAACTCAAGCCTTGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
|||||
2641 CCTCCAGTTCGTCGCCATCACAGCTTATCCAGTACCTGAGAGCTGCCATCTTCTCTCT 2700
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
|||||
2701 GTGGCAATTTGTCCCATGCTCTC 2721
901 ValAlaPheValProCysLeu 907
|||||


```
QY 1081 TACACCTATTAGAGATTATCCACGTTTTCAGTCTGCCAAAGCTTTCAGAAATTCGAC 1140
Db |||||
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db |||||
QY 1141 CTAAGACATAATGAATCTACGAATTAAGATTGACACTTTCCAGCAGTCTGCTTAGCCTC 1200
Db |||||
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db |||||
QY 1201 CGATCGCTGAATTTGGCTTGGAGCAAAATTCGTATTATTCACCCCAATGCAATTTCCACT 1260
Db |||||
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db |||||
QY 1261 TTGCCATCCCTAAATAAGCTGACCTATCGCTCCAACTCTGCTGCTCTTTCCCTATAACT 1320
Db |||||
QY 421 LeuProSerLeuIleLysLeuAspLeuSerAsnLeuLeuSerPheProIleThr 440
Db |||||
QY 1321 GGGTTACATGGTTTAACTCAGCTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db |||||
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db |||||
QY 1381 TCATCTGAAAATTTCCAGAACTCAAGTTATAGAAATGCCTTATGCTTACAGAGTCTGT 1440
Db |||||
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db |||||
QY 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTTCTCAATCAATGGAATAAAGGTGACAA 1500
Db |||||
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db |||||
QY 1501 AGCAGTATGGAGCACTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Db |||||
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db |||||
QY 1561 GACCTTGAAAGATTCTCGCTTGACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db |||||
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db |||||
QY 1621 TGTTCCACTTCCAGGAGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
Db |||||
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleuIle 560
Db |||||
QY 1681 AGAATTTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db |||||
QY 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db |||||
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Db |||||
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db |||||
QY 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGGCTGGCTGGTGGATCGGTTCACT 1860
Db |||||
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db |||||
QY 1861 TTTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTGGCATGCTCATT 1920
Db |||||
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db |||||
QY 1921 GGTTTTGTGTCATTTTGTCTTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGACCCCTG 1980
Db |||||
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db |||||
QY 1981 GAGCTGGGGTCTCTGTGGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db |||||
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db |||||
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGACCATGGCCGAGTTCCTCCCTG 2100
Db |||||
QY 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db |||||
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTGGGAGGCC 2160
Db |||||
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
Db |||||
QY 2161 AGCACCATGGGTACATGGTGCCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
```

```
Db |||||
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db |||||
QY 2221 ACCATTTGCTACACCAAGCTCTACTGCAATTTTGACAAAGGAGAGCTGGAGAAATTTGG 2280
Db |||||
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db |||||
QY 2281 GACTGCTCTATGGTAAACACATTGGCCCTGTGCTTCCACCAACTGCATCTCTAAACTGC 2340
Db |||||
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db |||||
QY 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db |||||
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db |||||
QY 2401 ATTAAGTTTATCTTCTGCTGCTAGTCCCACTTCTGATGCTCAATCCCTCTCTCTAC 2460
Db |||||
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db |||||
QY 2461 ATCTTGTTCATCTCTCACTTTTAAGGAGGATCTGTGTAGCCTGAGAAAGCAAACTTACGTC 2520
Db |||||
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db |||||
QY 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATGTCAATTAACTCTGATGTCTGAAAAACAG 2580
Db |||||
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
Db |||||
QY 2581 TCCTGTGACTCACTCAAGCCCTTGGTAACTTTACCAGTCCAGCATCATTATGACCTG 2640
Db |||||
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
Db |||||
QY 2641 CCTCCCAGTTCGTCGCACTCACAGCTTATCCAGTACTGAGAGCTGCATCTTTCTCTCT 2700
Db |||||
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db |||||
QY 2701 GTGGCATTTGTCCCATGTCTC 2721
Db |||||
QY 901 ValAlaPheValProCysLeu 907
Db |||||

RESULT 5
ADC22783
ID ADC22783 standard; protein; 907 AA.
XX AC ADC22783;
XX DT 18-DEC-2003 (first entry)
XX DE Human G protein-coupled receptor (GPCR) polypeptide #39.
XX KW Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6; intracellular-3 region; IC3; receptor.
XX OS Homo sapiens.
XX PN US6555339-B1.
XX PD 29-APR-2003.
XX PF 13-OCT-1998; 98US-00170496.
XX PR 14-APR-1997; 97US-00839449.
XX PR 14-APR-1998; 98US-00060188.
XX PR 26-JUN-1998; 98US-0090783P.
XX PR 07-AUG-1998; 98US-0095677P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-742861/70.
XX DR N-PSDB; ADC22782.
XX PT Creating a constitutively active version of an endogenous human G protein
```


301 GCTCTGACATATCCCAAGGAGGACATTCACATGGCCCTTTACAGTCTTAAAGTCTCTATG 360
101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
361 CTCGAGAATAATCAGCTTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTTGCCAAGCCTT 420
121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
421 CAATCCCTGGCTGTGAATGTAAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
481 CTCGATTCCTGAGGACCTGCTGGTGAATGCAATGCGTTTACAGAAATCCCGCTCCAG 540
161 LeuHisSerLeuArgHisIleuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
541 GCTTTTAGAAGTTTATCGGATTCGCAAGCCATCAGCTTGGCCCTGGAACAAATACACCCAC 600
181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisIleuHisAsnAsn 220
661 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
721 TTAATAATACAACTTATGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTAAA 780
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
781 GAACTAGGATTTTCATAGCAACAATATCAGTGCATACCTGAGAAAGCATTTGAGCAAC 840
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
841 CCTTCTCTTATTACAATACATTTCTATGACAATCCCATCCCAATTTGTTGGAGACTGCT 900
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
901 TTTCAACATTTACTGAACACTAAGAACACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 960
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTAACTGGAGACAGATC 1020
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCAAGTGTAGATCTGTCT 1080
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1081 TACAACCTATTAGAAGATTTACCAGTTTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
1141 CTAAGACATAAATGAATCTACGAAATTTAAAGTTTGACACTTTCACGACAGTTGCTTAGCCCTC 1200
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
1201 CGATCCGCTGAATTTGGCTGGAAACAAAATTTGCTATTATTATCCACCCCAATGCTTTCCACT 1260
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTGTCGCTTTTCTCTATAACT 1320
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
1321 GGGTTACATGGTTTAACTCACTTAAATAATTAACAGGAAATCATGCCCTTACAGACTTGATA 1380
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCGCTTATGCTTACCAGTGTCTGT 1440

461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
1441 GCATTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGAATAAAGGTGCAAC 1500
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1501 AGCAGTATGCAGCAGCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1561 GACCTTGAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
1621 TGTTCACCTTCCCCAGGCGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGCTGGCTGATC 1680
541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTGTAATGCTTTGGTGACTTCA 1740
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCATTAATACTGTTAATTTGGGTCAATCGCA 1800
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
1801 GCAGTGAAACATGCTACGCGGAGTCTCCAGTGGCGTCTGCTGGCTGTGTGATGGTTCACCT 1860
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1861 TTTGGCAGCTTTGCACGACATGGTGGCTGGTGGGAGAAATGGGGTTGGTCCCATGTCATT 1920
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1921 GCTTTTTCCTCATTTTCTCTCAGAAATCATCTGTTTCTCTGCTTACTCTGCGCAGCCCTG 1980
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
1981 GAGCGTGGCTTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTTCATTTTCTAGC 2040
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
2041 CTGAAAGTAATCATTTTGTCTCTGTGCCCTGTGCTGCTGACCATGCGCAGTTCCTCCTG 2100
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
2101 CTGGGTGGCAGCAAGTATGCGGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
2161 AGCACCATGGCTTACATGCTGCTCTCATCTGCTCAATTCCTTTCCTTCTCATGATG 2220
721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuSerLeuCysPheLeuMetMet 740
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTTGG 2280
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
2281 GACTGCTCTATGGTAAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
2341 CTGTGCTGCTTCTGTGCTTCTCTCTTAAATAAACCTTACATTTATAGTCTGCTGAAGTA 2400
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
2401 ATTAAGTTTATCTTCTGTGTGTAGTCCCACTTCTGTCATGTCTCAATCCCTTCTCTAC 2460
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
2461 ATCTGTGTTCAATCTCACCTTTAAGGAGGATCTGGTGAGGCTGAGAAAGCAACCTACGTC 2520

QY 1021 TGATCTCTTCTCAAAACCGTCTGCAATCAGTACCTTAATCTCAAGTGTAGATCTGTCT 1080
DB 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAAGATTACCAGATTTTTCAGTCTGCGCAAAAGCTTCAGAAAATTGAC 1140
DB 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGCATAATGAATCTACGAAATTAAGTTGACACATTTCCAGAGTTGCTTACGCTC 1200
DB 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCCCTGAATTTGCTGGCAAAATTTGCTATTATTCACCCCAATGCTATTTCCACT 1260
DB 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTGCTTTTCTATAACT 1320
DB 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1380
DB 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCGAGTCTGT 1440
DB 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGAGAAATGCTATAGATTTCTAATCAATGGAATAAAGTGACAAC 1500
DB 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGAGACCTTCATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
DB 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCTGCTGCTTGTAGACTTTGAGGAAGACCTGAAAGCCCTTCATTGAG 1620
DB 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTTCCCGAGCCCTTCAACCCCTGTGAACCTGCTGATCGCTGCTGCTGATC 1680
DB 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTTGGAGTGTGGACCATAGCAGTCTGGCACTTACTTGTAAATGCTTTGGTGACTCA 1740
DB 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGTTAAATGGGTCATCGCA 1800
DB 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATCTCACGGAGCTTCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGAGAATGGGCTGCTGCTGCTGCTGCTGCT 1920
DB 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTATTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGCTTCTCTGTAATATTTCTGCAAAATTTGAACGAAAGCTTCATTTCTAGC 2040
DB 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 681 LeuLysValIleLeuLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160

DB 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGCTTACATGCTCGCTCTCAATTCCTTCAATTCCTTTCCTTTCCTCATGATG 2220
DB 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACATTTGGCTTACACCAAGCTCTACGTCAATTTGGACAAGGAGACCTGGAGAAATTTGG 2280
DB 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATCGTAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
DB 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CTTGCTGCTTCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
DB 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
DB 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2520
DB 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAGATCAAAACACCAAGCTTGTGATGCTCAATTAATCTTCTGATGCTGCAAAACAG 2580
DB 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CTCTCCAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
DB 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGCTCTC 2721
DB 901 ValAlaPheValProCysLeu 907
RESULT 9
ADH14256
ID ADH14256 standard; protein; 907 AA.
XX ADH14256;
AC ADH14256;
XX 11-MAR-2004 (first entry)
DT 11-MAR-2004 (first entry)
XX Human HG38.
DE
XX human; non-endogenous; G protein-coupled receptor; GPCR; receptor.
XX Homo sapiens.
XX US2003105292-Al.
XX 05-JUN-2003.
XX 20-SEP-2002; 2002US-00251385.
XX 26-JUN-1998; 98US-0090783P.
XX 07-AUG-1998; 98US-0095677P.
XX 13-OCT-1998; 98US-00170496.
XX (LIAM/) LIAM C W.
XX (BEHA/) BEHAN D P.
XX (CHAL/) CHALMERS D T.
PI Liam CW, Behan DP, Chalmers DT;
XX

Db 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Qy 121 GAGCCGAGGAGAGATGTTGCTCAGGGTGAAGTCTCCGACCTGGGCTCTCGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Qy 181 CCTTCCAACCTCAGCGTCTTCACTCCCTACCTTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTCGGAGAGATTACGCTCTTCCGGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
Qy 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCTTTACGTTCTAAAGTTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGATAATCAGCTAAGACACCTTACCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTCGCTCTGGATCTGAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATTTCCCTGAGGACCTGTGGCTGGATGACATGGGTTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTAGAAGTTTATCGCATTTGCAAGCATGACCTTGGCCCTGCAACAAAATACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Qy 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCTCCATCAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Qy 661 AGAATCCATCTCCTGGGAAAGAAATGCTTTGATGGGCTCCAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAATTAACAATAACTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAACTAGGATTTTCATAGCAACAATATCAGGTCGATFACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGACAATCCATCCCAATTTGTTGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCAACAATACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTCTGATTTAACTGGAACTGCAACCTCGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCGATTACCTTAATCTCCAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAAATCTACGAAATTAAGTTTGACACTTTCCAGCAGTGTCTTACGCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400

Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTCGTATTATTACCCCAATGTCATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaIlePheAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCGATCCCTTAATAAAGCTGAGCTATCGTCCAACCTCTCTGTCTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerAsnLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAACCTTTCCAGAACTCAAGCTTATGAAATGCCTTATGCTTACAGTGTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaIleGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGTGAGAATGCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCATATGAGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCTGCTTGTACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTACACCTTCCCAGGCCCCCTTCAAAACCTGTGAACACCTGCTGTGATGGTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTTGAGTGTGGAACCATAGCAGATTCTGCACCTTACTTGTATGCTTTTGGTGCATTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 CAGGTGAACATGCTCAGGGAGTCTCCAGTCCCGTGTGCTGTGTGGATGGCTGCTACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGACGACATGGTCCCTGGTGGGAGATGGGTGGTGGCCATGCTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGTCATTTTGTCTTCAAGATCATCTGTTTTTCTGTCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGCTGCAATGACATGGCGCGAGTCCCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTCTGCTTTCCTTTTGGGAGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGGCTACATGGTCCGCTCTCATCTGTGCTCAATCCCTTTGCTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTCGAGAATATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 121 LeuGlnAenAenGlnLeuAargHisValProThrGluAlaLeuGlnAenLeuAargSerLeu 140
Qy 421 CAATCCCTGCGCTGAGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuAargLeuAaspAlaAenHisLeSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATTCCCTGAGGCACTGTGGCTGGATGACAAATGGTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuAargHisLeuTrpLeuAaspAenAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTACAAGTTTATCGSCATTGCAAGCCATGACCTTGGCCCTGGAACAAATAATACACCAC 600
Db 181 AlaPheAargSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAenLysIleHisHis 200
Qy 601 ATACCAGACTATGCTCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCCATAACAAT 660
Db 201 IleProAaspTyrAlaPheGlyAenLeuSerSerLeuValValLeuHisLeuHisAenAen 220
Qy 661 AGAATCCACTCCTCGGGAAGAATGCTTTGATGGGCTCCAGCTCAGACCTAGACACTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAaspGlyLeuHisSerLeuGluThrLeuAasp 240
Qy 721 TTAAATTAACAATCACTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA 780
Db 241 LeuAenTyrAenAenLeuAaspGluPheProThrAlaIleAargThrLeuSerAenLeuLys 260
Qy 781 GAACTAGGATTTTCATAGCAACAATATCAGTGCATACCTCGAAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAenAenIleAargSerIleProGluLysAlaPheValGlyAen 280
Qy 841 CTTTCTCTTATTACAAATCATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAaspAenProIleGlnPheValGlyAargSerAla 300
Qy 901 TTTTCAACATTACCTGAAGTAAAGACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuAargThrLeuThrLeuAenGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTTCTGATTTAACTGGAACTGCAAAACCTGGAGACTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAaspLeuThrGlyThrAlaAenLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACTTAATCTCCAAGTCTAGATCTGCTCT 1080
Db 341 SerSerLeuProGlnThrValCysAenGlnLeuProAenLeuGlnValLeuAaspLeuSer 360
Qy 1081 TACAACTATTAGAAGATTTACCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAenLeuLeuGluAaspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAasp 380
Qy 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Db 381 LeuAargHisAenGluIleTyrGluIleLysValAaspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTGGAAACAAAATTGCTATTATTATCCACCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAenLeuAlaTrpAenLysIleAlaIleHisProAenAlaPheSerThr 420
Qy 1261 TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAaspLeuSerSerAenLeuLeuSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCTCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAenHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAACTTTCCAGAACTCAAGTTTATAGAAATGCCCTTATGCTTACCAGTCTGT 1440
Db 461 SerSerGluAenPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500

Db 481 AlaPheGlyValCysGluAenAlaTyrLysIleSerAenGlnTrpAenLysGlyAaspAen 500
Qy 1501 AGCAGTATGACGACCTTTCATAAGAAAATGCTGGAATGTTTCAGGCTCAGAGTACAGT 1560
Db 501 SerSerMetAaspAenLeuHisLysLysAaspAlaGlyMetPheGlnAlaGlnAaspGluAarg 520
Qy 1561 GACCTTCGAAGATTTCCCTGCTTACTTTGAGGAACACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAaspPheLeuLeuAaspPheGluGluAaspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCAGAGCCCTTCAAAACCTGTGAAACACCTGTGATGGCTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAaspGlyTrpLeuIle 560
Qy 1681 AGAATGGAGTGTGGACCATAGCAGTTCGGCACTTACTTGTATGCTTTGGTGCATCTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAenAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTTAAATTTGGGGTTCATCGCA 1800
Db 581 ThrValPheAargSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCCGTGTGGCTGTGTGGATGCGTTCATC 1860
Db 601 AlaValAenMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAaspAlaPheThr 620
Qy 1861 TTTTGGCAGCTTTGCACGACATGTCCTGTGGGGAGAAATGGGTTGGTGGCCATGTCATT 1920
Db 621 PheGlySerPheAlaAargHisGlyAlaTrpTrpGluAenGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTCGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGGTGGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGTCCATTTCTTAGC 2040
Db 661 GluAargGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAAATCATTTTGTCTGTCCTGTCCTGTGGCCTGACCATGGCCGAGTTCCTCCCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGTGGCAGCAAGATAGGCGCTCCCTCTCTGCTGCTGCTTTGCCCTTTTGGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGGCTACATGGTCGCTCATCTGTGCTCAATCCCTTTGCTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAenSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAenLeuAaspLysGlyAaspLeuGluAenIleTrp 760
Qy 2281 GACTGCTCTATGGTAAAAACACATTCCTGCTGCTTCTTCCACCAACTGATCTCTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAenCysIleLeuAenCys 780
Qy 2341 CCTGTGGCTTTCTGTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCGCTGAATA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAenLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAenProLeuLeuTyr 820
Qy 2461 ATCTGTTTCAATCTCACTTTAAGGAGGATCTGGTGAGCTCTGAGAGCTGAGAAAGCAACCTACGCTC 2520
Db 821 IleLeuPheAenProHisPheLysGluAaspLeuValSerLeuAargLysGlnThrTyrVal 840
Qy 2521 TGGACAGATCAAAACACCCAGCTTGATGCTCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAaspValGluLysGln 860

181 AlapheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIleHisHis 200
601 ATACAGACTATGCTTTGGAAACCTCCAGCTTGGTAGTTCTACACTCCATAACAAT 660
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGACACTTAGAT 720
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
721 TTAATAACAATAACTTGTAGTAATCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
781 GAACTAGGATTTCTATAGCAACAATATCAGGTGCTACCTGAGAAAGCAATTTGTAGGCAAC 840
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
841 CCTTCTCTTATTACATACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
901 TTTCACAATTTACCTGAACTAGAACACTGACTCTGAATGGTGCCTCACAATAAATCA 960
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
961 TTTCCTGATTTAACTGGAACCTGCAACCTGGAGACTCTGACTTTAACTGGAGCACAGATC 1020
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle 340
1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTATCTCCAAGTCTAGATCTGCT 1080
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1081 TACAACCTATTAGAGAGTTTACCAGTGTTCAGTCTGCCAAAGCTTCAGAAAATTCAC 1140
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTTAGCTC 1200
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTCACCCCAATGCCATTTCCACT 1260
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
1261 TTGCCATCTCTAATAAGCTGACCTATCGTCCAACCTCTGCTGCTTTTCTCTATAACT 1320
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCAATGCTTACAGAGCTTGATA 1380
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1381 TCATCTGAAAACCTTCCAGAACTCAAGTTATAGAAATGCCCTTATGCCATTACAGTCTGT 1440
461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
1441 GCATTTGGAGTGTGTGAGAAATCCCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAA 1500
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1501 AGCAGTATGACGACCTTCATAGAAGAAGTGTGGAATGTTTCAGGCTCAAGATGAACGT 1560
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1561 GACCTTGAAGATTTCTGCTTGCATTTGAGGAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGCTGAACACCTGCTTGATGGCTGGCTGATC 1680

541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1681 AGAATTGAGTGTGGACCATAGAGATTCTGCACCTTACTTGTAACTGTTGGTGCATTC 1740
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTTAAATGGGGCTCATCGCA 1800
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
1801 GCACTGAACATGCTCACGGGAGTCTCCAGTGCCCTGTGGCTGTGTGGATGCGTCTACT 1860
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1861 TTTCGAGCTTTGACACGACATGGTCCCTGGTGGGAGAAATGGGTTGGTGGCCATGTCATT 1920
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1921 GGTTTTTTGTCCATTTTGTCTTCAGAATCATCTGTTTCTCTGCTTACTCTCGCAGCCCTG 1980
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
1981 GAGCGTGGGTTCTCTGTCAAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
2041 CTGAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
2101 CTGGTGGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
701 LeuGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
2161 AGCACCATGGCTACATGTCGCTCATCTGCTCATCTGCTCATCTGCTGCTGCTGCTGCTG 2220
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGAGCTCGAGAAATTTGG 2280
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
2281 GACTGCTCTATGGTAAACACATTTGCCCTGCTGCTTTCACCAACATGATCCTTAACTGC 2340
761 AspCysSerMetValLysHisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCys 780
2341 CCTGTGGCTTTCTGTCTCTCTCTTTTAAACCTTACATTTATCAGTCTCCTGAAGTA 2400
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
2401 ATTAAGTTTATCCTTCTGCTGCTAGTCCACTTCTGCTGCTGCTCAATCCCTCTCTTAC 2460
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
2461 ATCTTGTTCATCTCACTTTAAGGAGGATCTGTGTAGCCTCGAGAAACCAACCTACGTC 2520
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuLeuGlyGlnThrTyrVal 840
2521 TGGACAAGATCAAAACACCCAGCTTGATGTAATTAATCTGTGATGATGTCGAAAAACAG 2580
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
2581 TCCTGTGACTCAACTCAAGCCTTGTGTACCTTTACCAGTCCAGCATCATTATCACCTG 2640
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
2641 CCTCCAGTTCCTGCCATCAACAGCTTATCCAGTGTGAGAGCTGAGAGCTGCTCTCTCT 2700
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
2701 GTGGCATTTGCTCCCATGCTC 2721
901 ValAlaPheValProCysLeu 907

QY 1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTACACCCCAATGCATTTTCCACT 1260
DB 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGGCATCCCTAATAAAGCTGGACCTATCGTCAACCTCCTCTGCTGCTTTTCTCTATAACT 1320
DB 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGCTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
DB 441 GlyLeuHisGlyLeuThrHisLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAAATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACACGAGCTGCT 1440
DB 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGAGAAATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
DB 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACGACCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGACGT 1560
DB 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAGATTTCTGCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGTCAG 1620
DB 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCGAGCCCTTCAAAACCTGTGAAACCTGTGATGCTGCTGCTGCTGATC 1680
DB 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTTGGAGTGGACCATAGCAGTTCTGGCACTTCTGCTTAAATGCTTTGGTGACTTCA 1740
DB 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAATGCGGGTCATCGCA 1800
DB 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 1920
DB 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
DB 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGGTCTCTGCTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAACTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
QY 2161 AGCACCATGGCTACATGGTCTGCTCATCTTCTCAATTCCTTCTGCTTCTCTCATGATG 2220
DB 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuLeuMetMet 740
QY 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGACAGGGGAGCACTGGAGAAATATTTGG 2280
DB 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGCTGAACACATTTGCCCTTGTCTCTTCAACCACTGCATCTCTAAACTGC 2340

DB 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGGGCTTTCTGCTCCTTCTCCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
DB 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
DB 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTGTTTCAATCTCCTCACTTTAAGGAGGATCTGGTGAGCTGAGAAAGCAACCTACGTC 2520
DB 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAGATCAAAACACCAAGCTTGTGATCAATTAATCTCTGATGATGATGATGATGATGATG 2580
DB 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCCTTGTGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCGCTGCCATCACAGCTTATCCAGTGAAGCTGAGAGCTGCCATCTTCTCTCT 2700
DB 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGCTCCATGCTC 2721
DB 901 ValAlaPheValProCysLeu 907
RESULT 15
ADR67868
ID ADR67868 standard; protein; 907 AA.
XX
AC ADR67868;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human HG38 protein.
XX
KW human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
XX
OS Homo sapiens.
XX
PN WO2004074436-A2.
XX
PD 02-SEP-2004.
XX
PF 11-FEB-2004; 2004WO-US004060.
XX
PR 19-FEB-2003; 2003US-0448959P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Lasek AW;
XX
DR WPI; 2004-652946/63.
DR N-PSDB; ADR67869.
XX
PT Detecting colon or lung cancer, by determining amount of protein in
PT sample, comparing amount of protein to standard, and differential
PT expression of protein in sample indicates colon or lung cancer.
XX
PS Claim 4; SEQ ID NO 1; 79pp; English.
XX
CC This sequence represents the human G protein-coupled receptor (GPCR)
CC known as HG38. The HG38 protein and corresponding nucleic acid, may be
CC used in the method of the invention for detecting colon or lung cancer.
CC The first method involves performing an assay to determine the amount of
CC HG38 in a sample of colon or lung tissue, and comparing the amount of
CC protein to standard, thus detecting expression of protein in sample,
CC where differential expression of protein in sample when compared with the


```
QY 1861 TTTGGCAGCTTTGACGACATGCTGCTGGTGGAGANTGGGTGGTTGGTCATGTCATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTGTGTCATTTTGTCTTCAAGATCATCTGTTTTCCTGCTTACTCTGCGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGGTTCTCTGTGAATATTCTGCAAAATTTGAAAGCTCCATTTTCTTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTCTCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCGCTGCTGCGCTGCTGCGCTGCTGCG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
QY 2161 AGCACCATGGGCTACATGGTCTCTCATCTTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGAGACCTGGAGAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTCGCCCTGTTGCTCTTCCAACTGCATGCTTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCCTTCTGCTGCTAGTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGGTGAGCTGAGAAAGCAAACTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAGATCAAAACACCCAGCTTGATGTCATTAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCACTCAAGCTTGTGTAACCTTTACAGCTCCAGCATCATTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTCCGTCGCCATCACAGCTTATCCAGTGACTGAGAGCTGCGCTCTTCTCTCTCT 2700
Db 881 ProProSerSerValProSerProLafyProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGCTC 2721
Db 901 ValAlaPheValProCysLeu 907
```

Search completed: July 13, 2005, 03:49:43
Job time : 375.5 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	907	100.0	907	4	US-09-170-496D-264	Sequence 264, App
2	806	88.9	907	4	US-09-170-496D-278	Sequence 278, App
3	9	1.0	131	4	US-09-252-991A-17184	Sequence 17184, A
c	9	1.0	135	4	US-09-252-991A-18526	Sequence 18526, A
c	9	1.0	268	4	US-09-252-991A-17449	Sequence 17449, A
6	9	1.0	644	2	US-08-866-757-2	Sequence 2, Appli
7	9	1.0	644	3	US-09-153-593-2	Sequence 2, Appli
8	9	1.0	1133	4	US-10-101-464A-809	Sequence 809, App
9	8	0.9	99	4	US-09-482-273-258	Sequence 258, App
10	8	0.9	118	4	US-09-252-991A-25344	Sequence 25344, A
c	11	8	149	4	US-09-252-991A-24978	Sequence 24978, A
c	12	8	157	4	US-09-252-991A-28441	Sequence 28441, A

Db 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
QY 121 GAGCCGCGAGGATGTTGCTCAGGGTGGATGCTCCGACCTGGGCTCTCGGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
QY 181 CTTTCAACCTCAGGCTTTCACCTCCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
QY 241 CTGCTCCCGAAATCCCTCGCCTCCGCTTCTCGGAGGATGACGCTCTCGCGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
QY 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
QY 361 CTGCAGATATACGCTAAGACACGCTACCCACAGAGCTCTGCAGAAATTTGGAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
QY 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
QY 481 CTGCATTCCTGAGCACCTGCTGGTGGATGACAAATGCGTTAAACAGAAATCCCTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
QY 541 GCTTTTGAAGTTTATCGCATTCGACATTCAGCCATGACCTTGGCCCTGAAACAATAACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
QY 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
QY 661 AGAATCCATCTCCTGGGAAGAAATGCTTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
QY 721 TTAATTTACAATAACCTTGATGAATTTCCCACTGCAATTAGCACACTCTCCACCTTAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
QY 781 GAATAGGATTTTCATAGCAACAATATCAGTCTGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
QY 841 CTTTCTCTTATTAACAATATTTATGCAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTTACCTGAACCTAGAACACTGACTCTGAATGGTGCCTCACAAATACCTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTTCTGATTTAACTGGAACTGCAACACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle 340
QY 1021 TCATCTCTTCCCAACCGCTCGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAGATTTACCGATTTTTCAGTCTGCGCAAAAGCTTCAGAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTGTGCTTACGCTC 1200

Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTATTACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCGCATCCCTAAATAAGCTCGACCTATCGTCCAAACCTCCCTGCTGCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGGTTTAACTCAGCTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCCAGAACTCAGGTTATAGAAATGCCTTATGCTTACAGGCTGCT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAATGCCTATAAGATTTCTTAATCAATGAAATAAAAGGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACGCTTTCATAAGAAAGATGCTCGAAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCCCTGCTTACCTTGGAGAAAGCCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluLeuAspPheGluLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTCTCCCAAGGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGCATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATGGAGTGTGGACCATAGCAGTCTCGACACTTACTTCTTAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAAATTTGGGGCTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGTGGTGTGGATCGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGACGCTTTGACGACATGGTGCCTGTGGGAGAAATGGGGTGGTGGTCCCATGTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTGTTCATTTTGTCTCAGAACTATCTGTTTTCCTGCTTACTCTGCGAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATGCTCTACCAAGCTCTTCTGCAATTTGGCAATGGGAGACTCGAGATATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAAACCTCCTGTCGTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAAACCTTCCAGAACTCAAGGTTATAGAAATCCCTTATAGCTTACCAGTGTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGAGTGTGTGAGAATGCCCTATTAAGATTTCTAATCAATGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCACTATGGAGCCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCTGCTGACCTTTCAGGAAGACCTGAAAGCCCTTCATTCACTGTCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGTGAACACCTGCTTGTAGTGGCTGTATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTTGAGTGTGGACCATAGCAGTCTGGCAGCTTACTGTGAATGCTTTGGTGAATTC 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCACTTAAACTGTTAAATGGGGTCATGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTCCGCTGGCTGGTGGTGGTGGTGGTGGTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGCACGACATGGTCCCTGGTGGGAGAATGGGTGGTGGTGGTGGTGGTCA 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAGTAATCATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCGACGACATGATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPhePheGlyGluPro 720
QY 2161 AGCACCATGGGTACATGGTGGCTCTCATCTGCTCAATTCCTTGGCTTCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGATAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGT-AAAACACATTTGCCCTGTGCTTCTTCCAACTGCATCTCAAACTG 2339
Db 761 AspCysSerMetLysLys-HisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCy 780

QY 2340 CCCTGTGCCTTTCTTGTCTTCTCCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGT 2399
Db 780 sProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
QY 2400 AATTAAGTTTATCCTTTCTGGTGGTAGTCCCACTTCTTCGCAATGTCTCAATCCCTTCTCTA 2459
Db 800 ILeLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
QY 2460 CATCTTGTCAATCCTCATTAAAGGAGATCTCGTGAGCTCGAGAAAGCAAACTACGT 2519
Db 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVa 840
QY 2520 CTGGACAGATCAAAACACCCCAAGCTTGATCTCAATTAACCTCTGATGATGTCGAAAAACA 2579
Db 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2580 GTCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACCAGCTCCAGCATCACTTATGACCT 2639
Db 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLe 880
QY 2640 GCCTCCAGTTCCTGGCCATCCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTC 2699
Db 880 uProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSe 900
QY 2700 TGTGCGCATTTGTCCTCATCTCTC 2721
Db 900 rValAlaPheValProCysLeu 907

RESULT 3
US-09-252-991A-17184
; Sequence 17184, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17184
; LENGTH: 131
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17184

Alignment Scores:
Pred. No.: 10.8 Length: 131
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-17184 (1-131)
QY 64 CCCCCTGGCCAGCTGCAGCAGCACAG 38
Db 63 ProArgSerProAlaAlaAlaGln 71

RESULT 4
US-09-252-991A-18526
; Sequence 18526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18526
LENGTH: 135
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18526

Alignment Scores:
Pred. No.: 10.7 Length: 135
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-18526 (1-135)

QY 56 CGACCGGGGCGACCTCCCGTCTG 82
DB 116 ArgProGlyAlaAlaLeuProGlyLeu 124

RESULT 5

US-09-252-991A-17449
Sequence 17449, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17449
LENGTH: 268
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17449

Alignment Scores:
Pred. No.: 9.65 Length: 268
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-17449 (1-268)

QY 63 CCGGTCCGCGACGTGACGACGAGG 37
DB 160 ProGlyArgGlnLeuGlnGlnHisArg 168

RESULT 6

US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DERK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2

Alignment Scores:
Pred. No.: 8.43 Length: 644
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x US-08-866-757-2 (1-644)

QY 1420 CCTTATGCTTACCAGTGTGTCATT 1446
DB 158 ProTyrAlaTyrGlnCysCysAlaPhe 166

RESULT 7

US-09-153-593-2
Sequence 2, Application US/09153593A
Patent No. 6174994
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DERK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
FILE REFERENCE: GH-70055-1
CURRENT APPLICATION NUMBER: US/09/153,593A
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/866,757
EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 644
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-153-593-2

Alignment Scores:
Pred. No.: 8.43 Length: 644
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-153-593-2 (1-644)

Qy 1420 CCTATGCTTACCAAGTCTGTGCAATTT 1446
Db 158 ProTyrAlaTyrGlnCysCysAlaPhe 166

RESULT 8

US-10-101-464A-809
; Sequence 809, Application US/10101464A

; Patent No. 6768041
; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling

; FILE REFERENCE: 11000.1020C2

; CURRENT APPLICATION NUMBER: US/10/101,464A

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 809

; LENGTH: 1133

; TYPE: PRT

; ORGANISM: Eucalyptus grandis

US-10-101-464A-809

Alignment Scores:
Pred. No.: 7.73 Length: 1133
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-10-101-464A-809 (1-1133)

Qy 1060 CTCGAAGTCTAGATCTGTCTTACAAC 1086
Db 175 LeuGlnValLeuAspLeuSerTyrAsn 183

RESULT 9

US-09-482-273-258

; Sequence 258, Application US/09482273

; Patent No. 6534631

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 71 Human Secreted Proteins

; FILE REFERENCE: P2030P1

; CURRENT APPLICATION NUMBER: US/09/482,273

; CURRENT FILING DATE: 2000-01-13

; EARLIER APPLICATION NUMBER: PCT/US99/15849

; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,922

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,956

; EARLIER FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 258

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-482-273-258

Alignment Scores:
Pred. No.: 102 Length: 99
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-482-273-258 (1-99)

Qy 2230 TACCAAGCTCTACTGCAATTTG 2253

Db 33 TyrThrLysLeuTyrCysAsnLeu 40

RESULT 10

US-09-252-991A-25344

; Sequence 25344, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25344

; LENGTH: 118

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25344

Alignment Scores:
Pred. No.: 99.4 Length: 118
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-25344 (1-118)

Qy 42 GCTGCTGCAGCTGGCGACCGGGG 65

Db 33 AlaAlaAlaGlyAspArgGly 40

RESULT 11

US-09-252-991A-24978

; Sequence 24978, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24978
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24978

Alignment Scores:
Pred. No.: 95.9 Length: 149
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.92%
Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-24978 (1-149)

QY 84 ACCAGACTGGAGAGCTGCCCCC 61
Db 36 ThrA-gProGlyArgAlaAPro 43

RESULT 12

US-09-252-991A-28441
; Sequence 28441, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28441
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28441

Alignment Scores:
Pred. No.: 95.1 Length: 157
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.92%
Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-28441 (1-157)

QY 78 CTTGGAGAGCTGCCCGCTGCC 55
Db 40 ProGlyArgAlaAProGlyArg 47

RESULT 13

US-09-252-991A-28201
; Sequence 28201, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28201
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28201

Alignment Scores:
Pred. No.: 94.8 Length: 161
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.88%
Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-28201 (1-161)

QY 56 CGACCGGGGAGCTCTCCAGGT 79
Db 79 ArgProGlyAlaAAlaLeuProGly 86

RESULT 14

US-09-270-767-60345
; Sequence 60345, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-034
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 60345
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-60345

Alignment Scores:
Pred. No.: 91.5 Length: 203
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.88%
Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-270-767-60345 (1-203)

QY 486 TTCCCTGAGGCACCTGTGCTGGA 509
Db 68 PheProGluAlaProValAlaGly 75

RESULT 15

US-09-482-273-167
; Sequence 167, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-167

Alignment Scores:
Pred. No.:      91.2      Length:      207
Score:          8.00      Matches:      8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      0.88% Indels:      0
DB:              4      Gaps:      0

US-10-751-736-21 (1-2724) x US-09-482-273-167 (1-207)

Qy  2428 CCACCTTCCTGCATGCTCAATCCC 2451
    |||||||||||||||||||||
Db  47 ProLeuProAlaCysLeuAsnPro 54
```

Search completed: July 13, 2005, 04:00:05
Job time : 75 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: July 13, 2005, 03:40:32 ; Search time 194.5 Seconds
(without alignments)
10821.262 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 907

Sequence: 1 atggacacctccgctcgg.....catttgtccatgtctctaa 2724

Scoring table:

OLIGO 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1726218 seqs, 386331768 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3311304

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ n2p.model
-O=/cgn2_1/USPTO_spool/US10751736/runat_12072005_084330_23973/app_query.fasta_1.2887
-DB=Published Applications AA -QWTF=fastan -SUFFIX=oli.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USRP=US10751736 @CGN 1.1.278 @runat_12072005_084330_23973
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubaa/US10E_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubaa/US11A_PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB.pcp.*
21: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pcp.*
22: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pcp.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	907	100.0	907	10	US-09-965-536A-16	Sequence 16, Appl
2	907	100.0	907	10	US-03-970-944-26	Sequence 26, Appl
3	907	100.0	907	14	US-10-251-385-264	Sequence 264, App
4	907	100.0	907	14	US-10-225-567A-422	Sequence 422, App
5	907	100.0	907	15	US-10-295-027-849	Sequence 849, App
6	907	100.0	907	15	US-10-295-027-946	Sequence 946, App
7	907	100.0	907	15	US-10-295-027-1331	Sequence 1331, Ap
8	907	100.0	907	15	US-10-751-736-84	Sequence 84, Appl
9	907	100.0	907	17	US-10-482-029-158	Sequence 158, App
10	806	88.9	907	14	US-10-251-385-278	Sequence 278, App
11	806	88.9	1145	17	US-10-505-486-103	Sequence 103, App
12	716	78.9	907	10	US-09-970-944-27	Sequence 27, Appl
13	621	68.5	883	15	US-10-295-027-484	Sequence 484, App
14	621	68.5	883	15	US-10-295-027-1330	Sequence 1330, Ap
15	621	68.5	883	15	US-10-173-999-28	Sequence 28, Appl
16	592	65.3	693	17	US-10-851-470-4	Sequence 4, Appl
17	197	21.7	282	15	US-10-264-049-3132	Sequence 3132, Ap
18	51	5.6	907	14	US-10-271-078-10	Sequence 10, Appl
19	20	2.2	20	14	US-10-225-567A-1645	Sequence 1645, Ap
20	20	2.2	20	14	US-10-225-567A-1649	Sequence 1649, Ap
21	18	2.0	736	10	US-03-851-595-8	Sequence 8, Appl
22	18	2.0	736	15	US-10-664-667-8	Sequence 8, Appl
23	18	2.0	794	14	US-10-270-336-7	Sequence 7, Appl
24	18	2.0	823	14	US-10-270-336-6	Sequence 6, Appl
25	18	2.0	828	10	US-09-970-944-25	Sequence 25, Appl
26	18	2.0	828	14	US-10-270-336-2	Sequence 2, Appl
27	18	2.0	828	14	US-10-225-567A-581	Sequence 581, App
28	18	2.0	893	10	US-09-970-944-24	Sequence 24, Appl
29	18	2.0	915	14	US-10-270-336-5	Sequence 5, Appl
30	18	2.0	915	15	US-10-331-496A-68	Sequence 68, Appl
31	18	2.0	915	15	US-10-295-027-1336	Sequence 1336, Ap
32	18	2.0	928	16	US-10-783-528-101	Sequence 101, App
33	18	2.0	940	10	US-09-970-944-6	Sequence 6, Appl
34	18	2.0	967	10	US-09-851-595-2	Sequence 2, Appl
35	18	2.0	967	10	US-09-851-595-11	Sequence 11, Appl
36	18	2.0	967	15	US-10-331-496A-95	Sequence 95, Appl
37	18	2.0	967	15	US-10-684-667-2	Sequence 2, Appl
38	18	2.0	967	15	US-10-684-667-11	Sequence 11, Appl
39	18	2.0	967	15	US-10-684-667-11	Sequence 11, Appl
40	18	2.0	967	16	US-10-737-450-32	Sequence 32, Appl
41	18	2.0	1005	14	US-10-176-847-90	Sequence 90, Appl
42	18	2.0	1205	17	US-10-505-486-64	Sequence 64, Appl
43	17	1.9	120	14	US-10-276-340-2	Sequence 2, Appl
44	17	1.9	948	16	US-10-338-036-10	Sequence 10, Appl
45	16	1.8	16	14	US-10-225-567A-1646	Sequence 1646, Ap

ALIGNMENTS

RESULT 1
US-09-965-536A-16
; Sequence 16, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY5,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
; FILE REFERENCE: D0041NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19

QY 841 CCTTCTCTTATACAAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCACAAATTTACCTGAACACTGAACACTCTGAATGTGTCTCAAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
QY 1021 TCATCTCTTCTCAAAACCGTCTGCATCAGTACATCTCAATCTCCAAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAAATTTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAChAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGGTTAACTCAGTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCAGAACTCAAGTTATAGAAATGCCTTATGCTTACAGTGCCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACGCTTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGNAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCGAGGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTTGGAGTGTGGACCATAGCAGTCTTGCCACTTACTTGAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACCTGTTAACTTTGGGGTCAATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCAGGGAGTCTCCAGTGGCGTGGCTGGTGTGTGATGCTGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTCGACGCTTTCACGACATGCTGCTGGTGGGAGATGGGGTTGGTTGCCATGCTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640

QY 1921 GGTTTTTTGTCCATTTTTTGTTCAGAAATCATCTGTTTTCTGCTTACTCTCGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGTGTCTCTGTGAAATATTCTGCAAAATTTGAAAGAAAGTCTCAATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTTGCTCTGTGCGCTGTGCTGCTGACCATGGCGGAGTTCCCTCG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTTTTGGGGAGCCC 2160
Db 701 LeuGlyLysLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCAACCATGGCTACATGGTGTCTCATCTGTGTCAATTCCTTTGCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTCGCTACCAAGCTCTACTGCAATTTGCAAGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTGGCCCTGTGTCTTTCACCAACTGCATCTTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTTCCTTCTCTCTTCTTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTATCTCTTCTGCTGCTAGTCCACTTCTGCTGCTGCTCAATCCCTTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGAAGCTGAGAAAGCAACCTAGCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGCAAGATCAAAACACCCCAAGCTTGATGTCATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGATCAACTCAAGCTTGTAACTTACCAGTCTTACCAGTCCAGCATCACATTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCCTGTCATCACACAGCTTATCCAGTGTATCCAGTGTGAGAGTCCATCTTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGCTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 3
US-10-251-385-264
; Sequence 264, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294


```

|||||
641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGTCTCTGTGAATATCTCAAAATTTGAAACGAAAGCTCCATTTCTTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAGTAATCAATTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGGTACATGGTGGCTCTCATCTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGNAATATTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
QY 2281 GACTGCTCTATGGTAAACACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCGTGGCTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAGTTTNCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCAATCTCTCACTTTAAGGAGATCTGGTGAGCCTGAGAAACCAACCTACCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuLeuArgLysGlnThrTyrVal 840
QY 2521 TGGCAAGATCAAAACACCAAGCTTGATGCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGATCAACTCAAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCCCCAGTTCCGTCCTCATCCAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCTCATGCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 4
US-10-225-567A-422
; Sequence 422, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 422

```

```

; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-422
Alignment Scores:
Pred. No.: 0 Length: 907
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-751-736-21 (1-2724) x US-10-225-567A-422 (1-907)
QY 1 ATGAGACACTCCCGGCTCGGTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
QY 61 GGGGGAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTGTCATTGC 120
Db 21 GlyGlySerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
QY 121 GAGCCCGACGCGAGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
QY 181 CCTTCGAACCTCAGCGTCTTCACCTCTACTAGACCTCAGTATGAACACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
QY 241 CTGCTCCCGAATCCCTGCGGCTCTCGGCTCTCTGAGGAGTTCACGCTTTCGCGGGAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuLeuArgGlyAsn 100
QY 301 GCTCTGACATACATTCCTCAAGGAGCATTCACCTGCGCTTACAGCTTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
QY 361 CTGAGATAATCAGCTAAGACACGCTACCCACAGACCTCTGCAGAAATTTGCGAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
QY 421 CAATCCCTGCTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
QY 481 CTGCATTCCTTGAGGCACTGTGCTGGATGACAATGCGTTTAAAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
QY 541 GCTTTTAGAAGTTTATCGGCATTCGAAGCCATGACCTTGGCCCTGACAAATATACCCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
QY 601 ATACCAGACTATGCTTTGGAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
QY 661 AGAATCCACTCCTCGGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
QY 721 TTAATTAACATAACTTGAATTCCTCCACTGCAATTAGAGACTCTCCAACTTAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
QY 781 GAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCTCGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
QY 841 CTTCTCTTATTACATACTTTCTATGACAATCCCATCCATTTTGTGGAGATCTGCT 900

```

Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTACTGAACACTAAGAACACTGACTCTGAATGGTGGCTCCACAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuAArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTGATTAACTGGAACCTGAAACCTGGAGAGTCTGACTTTAACTGAGCAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
QY 1021 TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAACTGCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TCAACCTATTGAAGATTACCACTTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAAATCAATCTACGAAATTAAGTTGACACTTCCAGCAGTGTGCTTACGCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCTCTCAATTTGGCTTGGACAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCTTAATAAGCTGACCTATCGTCAACCTCTCTGCTGCTTTTCTCTAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAATTTTCCAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTCT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTGGAGTGTGAGAAATGCTTAAAGATTTCTTAATCAATGAATAAAGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGACGACCTTCATAGAAGATGCTGGAATGTTCAAGCTCAAGATCAAGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTCTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGCGAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTTCCAGGCCCTTCAAAACCTGTGAAACCTGCTGACACCTGCTTGTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTGGAGTGTGACCATAGAGTCTGGCACTTCTGGAATGCTTGAATGCTTGTGATCTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAATCTTAATGGGCTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATCTCACGGAGCTCCAGTGGCGCTGCTGGTGTGATGCTGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGAGCTTTGACAGCATGCTGCTGGTGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTGTTTGTGCTTCTCAGAAATCATCTGTTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660

QY 1981 GAGCGTGGGTTCTCTGTGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACAATGGCTACATGCT 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTACCAACGCT 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGCTAAACACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CTTGTGGCTTCT 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCT 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrVal 840
QY 2521 TGGCAAGATCAAAACACCAAGCTTGTGATGCTCAATTAATCTCTGATGATGCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGATCACTCAAGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGAT 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CTTCCAGTTCGTCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 5

US-10-295-027-849
; Sequence 849, Application US/10295027
; Publication No.: US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13

Qy	1621	TGTTCACTTCCCAGGCCCTTCAAAACCCCTGTGAAACACCTGCTTGATGAGCTGGCTGATC	1680
Db	541	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
Qy	1681	AGAATTGGAGTGTGCACCATAGCAGTTCTGCACACTTACTTGTATGTCCTTTGGTGACTTCA	1740
Db	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
Qy	1741	ACAGTTTTCAGATCCCCTCTGTACATTTCCCCATTAAACTGTTAAATTGGGGTCATCGCA	1800
Db	581	ThrValPheArgSerProLeuTyriLeSerProIleLysLeuLeuIleGlyValIleAla	600
Qy	1801	GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGCCTGGTGGATGGTTCACT	1860
Db	601	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620
Qy	1861	TTTGGCAGCTTTGCAACGACATGGTGCTGGTGGGAGAAATGGGGTTGTTGCCATGTCAATT	1920
Db	621	PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle	640
Qy	1921	GGTTTTTTTGCCATTTTTTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	641	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu	660
Qy	1981	GAGCGTGGGTCTCTGTGAAATATTTCGCAAAATTTGAAACGAAAGCTCCATTTCTAGC	2040
Db	661	GluArgGlyPheSerValLysTySerAlaLysPheGluThrLysAlaProPheSerSer	680
Qy	2041	CTGAAAGTAATCATTTTGTCTGTGCCCCTGCTGGCCCTTGACCATGGCGCGAGTTCCCTCG	2100
Db	681	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetalaAlaValProLeu	700
Qy	2101	CTGGGTGGCAGCAAGTAGCGCCTCCCTCTCTGCGCTTGTGCTTTGGCTTTGGGGAGCCC	2160
Db	701	LeuGlyGlySerLysTyGlyAlaSerProLeuLysLeuProLeuProPheGlyGluPro	720
Qy	2161	AGCACCATGGCTACATGGTCGCTCATCTTGCTCAATTCCCTTTGCTTCCCTCATGATG	2220
Db	721	SerThrMetGlyTyMetValAlaLeuIleLeuLeuAsnSerLeuLysPheLeuMetMet	740
Qy	2221	ACCATTGCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAGATATTGG	2280
Db	741	ThrIleAlaTyThrLysLeuTyCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp	760
Qy	2281	GACTGCTCTATGTPAAACAACATGGCCCTGTGCTCTTCCACCACTGCATCTCTAAACTGC	2340
Db	761	AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
Qy	2341	CGTGTGGCTTCTGTGCTCTCTCTCTTTTAAACCTTAAACCTTAACTTATCATGCTCTGAGTA	2400
Db	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
Qy	2401	ATTAAGTTTATCCTTCTGGTGTAGTCCCACTTCCCTGCATGTCTCAATCCCTTCTCTAC	2460
Db	801	IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy	820
Qy	2461	ATCTTGTTCATCTCTACTTTAAGGAGATCTGTGTAGCCTTGAGAAAGCAAACTTAGCTC	2520
Db	821	IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyVal	840
Qy	2521	TGGACAAGATCAAAACCAACCAAGCTTCATGTCATTAACCTCATGATGATCGCAAAACAG	2580
Db	841	TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln	860
Qy	2581	TCCTGTGACTCAACTCAAGCCTTCGTTAACTTTTACCAGCTCCAGCATCACTTATGACCTG	2640
Db	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyAspLeu	880
Qy	2641	CCTCCCAGTTCCGTGCCATACACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2700
Db	881	ProProSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer	900

```

Qy      2701  GTGGCATTTTGTCCCATGTCTC 2721
      |||||
Db      901  ValAlaPheValProCysLeu 907

RESULT 6
US-10-295-027-946
; Sequence 946, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezai, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis
; TITLE OF INVENTION: Methods of Diagnosis
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,717
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,616
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,333
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,414
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,313
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,313
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,212
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,313
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,212
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,717
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 946
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-946

```

Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Qy 181 CCTTCCAACCTCAGGCTCTTCACTCCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAATCCCTCCAGTCCGCTCCGCTTCTCGGAGGAGTTCAGTCTTGGCGGAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
Qy 301 GCTCTGACATACATTTCCCAAGGAGCATTCAGTCCCTTTTACAGTCTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGATAATACGCTAAGACACGTACCCACAGAAAGCTCTGCAGAAATTCGGAAGCCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTCGCTCGATGCTAAACACATCAGCTATGTCCTCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCAATCCCTCAGGACACTGCTGGCTGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTAGAAGTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGCAACAATAATACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Qy 601 ATACAGACTATGCTTTGGAAACCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerLeuValLeuHisLeuHisAsnAsn 220
Qy 661 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTAGACACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAAATACATAACTTGATGAATTCCTCCACTGCAATTAGACACTCTCCAACTTAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAACTAGCATTTATAGCAACAATACAGTTCGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CCTTCTCTATTACATAATCTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTCTCTCAAAACCGCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTACCGATTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAATCTACGAATAATAAGTTGACACTTTTCCAGCAGTGTCTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420

Qy 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACCTCCTGTGCTCTTTTCTTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLysSerSerAsnLeuLeuSerPheProIleThr 440
Qy 1321 GGGTTACATGTTTAACTCACTTAAAAATTAAACAGAAATCATGCTTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAATCTTTCAGAACTCAAGGTTATAGAAATGCTTTATGCTTTACCAGTGTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGAGTGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAGGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAATTTCTGCTTGTGACTTTGAGAAAGACCTGAAAGCCCTTCATTCACTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCAGGCGCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTCGAGTGTGGACCATAGCAGTTCCTGGCACTTACTTGTATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGTCAATGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCTGCTGCTGTGTGGATGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGAGAAATGGGGTTGGTGCATGTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGCATTTTGTCTGCAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Qy 1981 GAGCGTGGTTCCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCCAATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAATCATTTTGTCTGTGCCCTGCTGCTGACCATGGCGGAGTTCCTCGT 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGCCCTTTTGGGGAGCCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGCTACATGGTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGCAAGGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACAACATTTGCCCTGTGTCTTCCACCACTGCATCTCAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780

QY 901 TTTCAACATTACCTGAACACTAAGACACTGACTCTGAATGCTGCTCACAATAACTCAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAenGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTGATTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAenLeuGluSerLeuThrLeuGlyAlaGlnIle 340
QY 1021 TCATCTCTTCTCAAAACCGTGTGCAATCAGTTACCTAACTCTCAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAenGlnLeuProAenLeuGlnValLeuAenLeuSer 360
QY 1081 TACAACCTATTAGAGATTACCCAGTTTTCAGTCTGCCAAAGACTTCAGAAAATGAC 1140
Db 361 TyrAenLeuLeuGluAenLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATAGAAATCTAGCAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db 381 LeuArgHisAenGluIleTyGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAenLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGACCTATCGTCCACCTCTGCTGCTTTTCTCTATACT 1320
Db 421 LeuProSerLeuIleLysLeuAenLeuSerSerAenLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAenHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db 461 SerSerGluAenPheProGluLeuLysValIleGluMetProTyAlaTyGlnCysCys 480
QY 1441 GCATTTGAGTGTGTGAGAACTGCTATAGATTCTTAATCAATGAATAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAenAlaTyLysLysIleSerAenGlnTrpAsnLysGlyAspAen 500
QY 1501 AGCAGTGGAGGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAAGATTCTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTGCTGAC 1620
Db 521 AspLeuGluAenPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTCTCCAGGCGCCCTTCAAAACCTGTGAACACCTGCTTGTGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAenAspGlyTrpLeuIle 560
QY 1681 AGAATTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAenAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGATTAATTTCCGCCATTAACCTGTTAATGGGGTCAATGCCA 1800
Db 581 ThrValPheArgSerProLeuTyIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CCAGTGAACATGCTCACGGAGTCTCCAGTGGCTGGCTGGTGGTGGTGGTGGTGGTGGT 1860
Db 601 AlaValAenMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGACGACATGGTCCCTGGTGGGAGAATGGGGTTGGTGGCAGTCAATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAenGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGAACTCACTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660

QY 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTySerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAGTATGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGCTACATGGTGGCTCTCATCTGTGCTCAATTTCCCTTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyMetValAlaLeuIleLeuAenSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGGAAGGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyThrLysLeuTyCysAenLeuAspLysGlyAspLeuGluAenIleTrp 760
QY 2281 GACTGCTCTATGGTAACACATTTGCCCTGTGCTCTTCCAACTGCATCTCTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAenCysIleLeuAenCys 780
QY 2341 CCTGTGGCTTTCTTGTCT 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAenLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAenProLeuLeuTy 820
QY 2461 ATCTGTTTCATCTCTCACTTTAAGGAGGATCTGTGAGCTGAGAGAAAGCAACCTACCTC 2520
Db 821 IleLeuPheAenProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyVal 840
QY 2521 TGGAACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTGTATGTGCGAAGAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTACAGCTCCAGCAGTCCAGTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyAspLeu 880
QY 2641 CCTCCAGTTCCTGGCCATCACCACTTATCCAGTGTGAGAGCTGCGATCTTCTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTGTGCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907
RESULT 8
US-10-751-736-84
; Sequence 84, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10/751,736
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 907
; TYPE: PRT

QY 1021 TCATCTCTTCCCTCAAAACCGTCTGCAATCATAGTTACCTAAATCTCCAAGTGCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAGATTATCCAGTCTTTCAGTCTGCGCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLeuGlnLeuLeuAsp 380
QY 1141 CTAAGACATAATGCAATCTACGAAATATAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluLeuLeuGluLeuValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCGCTCAATTTGGCTTGGAAACAAAATGCTATTATTATTCACCCCAATGCTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
QY 1261 TTGGCATCCCTAATAAGCTGACCTATCGTCCAACTCTCTGCTCTCTCTCTCTCTCTCT 1320
Db 421 LeuProSerLeuLeuLeuLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuLe 460
QY 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCys 480
QY 1441 GCATTGGAGTGTGAGAGATCCCTATAGATTTCTAATCAATGNAATAAAGTGACAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 ACAGTATGACGACGCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAGATTTCTGCTGCACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCGAGGCCCTTCAACCCCTGGAACACCTGCTGTGATGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuLe 560
QY 1681 AGAATTGGAGTGTGACCATAGCAGTTCTGGCACTTACTTGAATGCTTTGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTCAATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuLeuGlyValIleAla 600
QY 1801 CGAGTGAACATGCTACGGAGTCTCAAGTGGCTGCTGGCTGGTGGTGGATCGGTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTGGCAGCATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTGTCCATTTTCTTCCAGATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700

QY 2101 CTGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCTGCTGCTTTCGCTTTTGGGAGCCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACCATGGGTACATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuMet 740
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTTGGCAAGGGAGACTGGAGATAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAAAACACATTCCTGCTGCTTCCACCAACTGCATCTTAACCTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGGGCTTCT 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTGTTTCAATCTCTCACTTTTAAAGAGGATCTGCTGAGCTGAGAAAGCAAACTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrVal 840
QY 2521 TGGCAAGATCAAAACACCCAGCTTGATGCAATTAATCTGTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCGTCGCCATCACAGCTTATCCAGTACGAGCTGAGAGCTGCCATCTTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCTGCTGCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 10
US-10-251-385-278
; Sequence 278, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251.385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170.496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-278

Alignment Scores:
Pred. No.: 0 Length: 907
Score: 806.00 Matches: 906
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1

Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACTATTAGAGATTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAATCTACGAATTAAGTTCAGCTTTCCAGCAGTCTCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCCCTATTAAC 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAATTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGCTGCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGAGTGTGTGAGAAATCCCTAATAAGATTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGGACGCTTCATAAGAAAGATCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAAGATTTCTCTGCTGACTTTTCAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGAAACCTGCTTGTAGTGGCTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTGAGTGTGACCATAGCATTCTGGCAGCTTCTGTAATGCTTTGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTTAAATTTGGGTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACTGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGATCGGTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGCGACCTTTGCACGACATGCTCCCTGGTGGGAGATGGGTGGTGGTCCATGCTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGTCCATTTTGTCTCAGAACTCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCTGGGTCTCTGTGGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAGTAATCATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720

Qy 2161 AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGAT 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCAATTCCTACACCAAGCTCTACTGCAATTTTGGACAAAGGAGACCTGGAGAAATATTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCCACCAACTGATCCTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGTGGCTTCTGTCTCTCTCTCTTAAATAAACCTTATCAATTTATCAGTCCCTGAAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Qy 2461 ATCTTGTTCATCTCTCTTTAAGGAGATCTGTGAGCTGTGAGCTGTGAGAAACCACTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Qy 2521 TGGACAGATCAAAACACCCAGCTTGTATGATCTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Qy 2581 TCCTGTGACTCAACTCAAGCTTGTGTAACCTTTTACCAGCTCCAGCATCATTTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Qy 2641 CCTCCGAGTTCCTGGCCATCACAGCTTATCCAGTGTATCCAGTGTGAGCTGCCATCTTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Qy 2701 GTGGCATTTGTCCTCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 12
US-09-970-944-27
; Sequence 27, Application US/0970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR FILING DATE: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-27

Alignment Scores:
Pred. No.: 0 Length: 907
Score: 716.00 Matches: 816
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 78.94% Indels: 2
DB: 10 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-970-944-27 (1-907)

QY 271 TTCCTGGAGGAGTTAGCTTTCGGGAAAGCGCTCTGACATACATATCCCAAGGAGCATTC 330
 Db |||||
 91 PheLeuGluGluLeuArgLeuAlaGlyAsnAlaLeuThrTyrIleProIysGlyAlaPhe 110
 QY 331 ACTGGCCTTTACAGTCTTAAAGTTCTTATGCTGCAGATAAATCAGCTAAGACACGCTACCC 390
 Db |||||
 111 ThrGlyLeuTyrSerLeuIysValLeuMetLeuGlnAsnGlnLeuArgHisValPro 130
 QY 391 ACAGAGCTCTCAGAAATTCGAAGCTTCAATCCCTGCGTCTGAGTCTAACACACATC 450
 Db |||||
 131 ThrGluAlaLeuGlnAsnLeuArgSerLeuGlnSerLeuArgLeuAspAlaAsnHisIle 150
 QY 451 AGCTATGTCCTCCCAAGCTGTTTCAGTGGCTGCATTCCTCAGGACCTGTGGCTGGAT 510
 Db |||||
 151 SerTyrValProProSerCysPheSerGlyLeuHisSerLeuArgHisLeuTrpLeuAsp 170
 QY 511 GCAATAGCTTAAACAGAAATCCCGTCCAGGCTTTTAGAAGTTTATCGCATTTGCAAGCC 570
 Db |||||
 171 AspAsnAlaLeuThrGluIleProValGlnAlaPheArgSerLeuSerAlaLeuGlnAla 190
 QY 571 ATGACCTTGGCCCTGAACAAAATACACACATACCAGACTATGCTTTCGAAACCTCTCC 630
 Db |||||
 191 MetThrLeuAlaLeuAsnIysIleHisIleProAspTyrAlaPheGlyAsnLeuSer 210
 QY 631 AGCTTGGTA-GTTCTACATCTCCATAA CAATAGAATCCACTCCCTGGGAAAGAAATGCTT 689
 Db |||||
 211 Ser-TrpValValLeuHisLeuHisAsnAsnArgIleHisSerLeuGlyLysCysPh 230
 QY 690 TGATGGCTCCACAGCCTAGACACTTAGATTAAATACATAATACCTTGATGAATCCC 749
 Db |||||
 230 eAspGlyLeuHisSerLeuGluThrLeuAspLeuAsnTyrAsnAsnLeuAspGluPhePr 250
 QY 750 CACTGCAATTAGGACACTCTCAACCTTAAAGAACTAGGATTTCTAGCAACAATATCAG 809
 Db |||||
 250 oThrAlaIleArgThrLeuSerAsnLeuLysGluLeuGlyPheHisSerAsnAsnIleAr 270
 QY 810 GTCGATACCTGAGAAAGCATTTGTAGGCAACCTCTCTCTTATTAACAATACATTTCTATGA 869
 Db |||||
 270 gSerIleProGluLysAlaPheValGlyAsnProSerLeuIleThrIleHisPheTyrAs 290
 QY 870 CAATCCCATCAATTTGTTGGGAGATCTGCTTTTCAACATTTACCTGAACTAAGACACT 929
 Db |||||
 290 pAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGluLeuArgThrLe 310
 QY 930 GACTCTGAATGTGCTCACAAATAACTGAATTTCTGATTTAACTGGAACCTGCAACCT 989
 Db |||||
 310 uThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGlyThrAlaAsnLe 330
 QY 990 GGAGAGCTGTACTTTAACTGGAGCAGATCTCATCTCTTCTCAAAACCGTCTGCAATCA 1049
 Db |||||
 330 uGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnG1 350
 QY 1050 GTTACTAATCTCCAAGTGTAGATCTGCTTACACCTATTAGAGATTTTACCCAGTTT 1109
 Db |||||
 350 nLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPh 370
 QY 1110 TTCAGTCTGCCAAAAGCTTCAGAAAATTCAGCTAAGACATATGAATCAATCAAGAAATTA 1169
 Db |||||
 370 eSerValCysGlnLysLeuGlnLysIleAspLeuArgHisAsnGluIleTyrGluIleLys 390
 QY 1170 AGTTGACACTTTCCAGCAGTGTCTTAGCTCCGATCCGATCGCTGGAATTTGGCTTGGAAACAAAT 1229
 Db |||||
 390 sValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaTrpAsnLysI1 410
 QY 1230 TGCTATTATTCACCCCAATGATTTTCCACTTTGGCATCCCTTAATAAGCTGGACCTATC 1289
 Db |||||
 410 eAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleLysLeuAspLeuSe 430
 QY 1290 GTCCAACTCTCTGCTCTTTTCTTATAACTGGGTTTACATGGTTTAACTCACTTAAATTT 1349
 Db |||||
 430 rSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLe 450

QY 1350 AACAGAAATCATGCTTACAGAGCTTGATATCATCTGAAAACTTTTCAGAACTCAAGT 1409
 Db |||||
 450 uThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPheProGluLeuLysVa 470
 QY 1410 TATCAAAATGCCTTATGCTTACAGTGTGTCATTTGGAGTGTGTGAGAAATGCCTATAA 1469
 Db |||||
 470 IleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGluAsnAlaTyrLys 490
 QY 1470 GATTTCTTAATCAATCGAATAAAGGTGCAACACAGCTAGTAGGACGACCTTTTCATAAGAAAGA 1529
 Db |||||
 490 eIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspLeuHisLysLysAs 510
 QY 1530 TGCTGGAATGTTTCAGGCTCAAGATGAAGTCAAGTCAAGTTCAGATTTCTCGCTTGCATTGCA 1589
 Db |||||
 510 pAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeuLeuAspPheG1 530
 QY 1590 GGAACACCTGAAAGCCCTTTCATTCAGTGCAGTGTTCACCTCCCGAGGCCCTTCAAAACC 1649
 Db |||||
 530 uGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGlyProPheLysPr 550
 QY 1650 CTGTGAAACACCTGTGATGGCTGATCAGAAATTTGGAGTGTGGACCATAGCATTTCT 1709
 Db |||||
 550 oCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThrIleAlaValIe 570
 QY 1710 GGCATTACTTGTAAATGCTTTTGGTCACTTCAACAGTTCAGATCCCTCTGTATATTC 1769
 Db |||||
 570 uAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerProLeuTyrIleSe 590
 QY 1770 CCCCATTAAACTGTTAATTTGGGCTCATCGCAGCAGTGAACATGCTCAGGGAGTCTCCAG 1829
 Db |||||
 590 rProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThrGlyValSerSe 610
 QY 1830 TGCCGTGCTGCTGTGATGGCTTTCACATTTTGGCAGCTTTGACGACCATGGTGCCTG 1889
 Db |||||
 610 rAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyAlaTr 630
 QY 1890 GTGGAGAAATGGGTTGGTTCCTCATGTCATTTGGTTCATTTTGGTTCATTTTGGTTCAGATC 1949
 Db |||||
 630 pTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGluSe 650
 QY 1950 ATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGTTCCTGTAATATTTCTGCG 2009
 Db |||||
 650 rSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerValLysTyrSerAl 670
 QY 2010 AAAATTTGAAACGAAAGCTCCATTTTCTAGCCTCGAAAGTAAATCAATTTGCTCTGTGCCCT 2069
 Db |||||
 670 aLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeuCysAlaLe 690
 QY 2070 GCTGCTTGTGACCATGGCGGAGTTCCCTCTGCTGGTGGCAGCAAGTATGGCGCTCCCC 2129
 Db |||||
 690 uLeuAlaLeuThrMetAlaAlaValProLeuLeuGlySerLysTyrGlyAlaSerPr 710
 QY 2130 TCTCTGCTGCTTTCCTTTTGGGGCCCGACCATGGCTACATGGTTCGCTCAT 2189
 Db |||||
 710 oLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLeuI1 730
 QY 2190 CTGTGCTCAATTTCCCTTTTCTCTCATCATGACCATTCCTACCAAGCTCTACTGCAA 2249
 Db |||||
 730 eLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCysAs 750
 QY 2250 TTTGGAAGAAGGAGACCTGGAGATATTTGGGACCTGCTCTATGGTAAACACATGGCCCT 2309
 Db |||||
 750 nLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAlaLe 770
 QY 2310 GTTCTCTTCCCAACTGCATCCCTAAACCTGCTGGCTTCTTCTGCTCTCTCTCTCTT 2369
 Db |||||
 770 uLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSerPheSerSerLe 790
 QY 2370 AATAAACCTTACATTTTACCTCTGAAAGTAAATTAAGTTTATCTCTTCTGGTGGTAGTCCC 2429
 Db |||||
 790 uIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeuValValPr 810
 QY 2430 ACTTCTGTCATGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTCTCTTAAAGGAGA 2489

Db 810 oleuProAlaCysLeuAsnProLeuLeuTyrlleLeuPheAsnProHisPheLysGluAs 830
Qy 2490 TCTGGTGGCCCTGAGAAACCAACTACGCTCTGGACAAGATCAAAACACCCCAAGCTTGAT 2549
Db 830 pLeuValSerLeuArgLysGlnThrTyrlValTrpThrArgSerLysHisProSerLeuWe 850
Qy 2550 GTCATTAACCTCTGATGATGTCGAAAAACAGTCTCTGTGACTCAACTCAGGCTTGGTAAC 2609
Db 850 tSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGlnAlaLeuValTh 870
Qy 2610 CTTTACCAGCTCCAGCATCTTATGACCTCTCCCAAGTTCCGTCGCATCACCAGCTTA 2669
Db 870 rPheThrSerSerSerIleThrTyrlAspLeuProProSerSerValProSerProAlaTy 890
Qy 2670 TCCAGTGAAGTCCAGCTCCCACTTTCTCTGTGGCAATTTCTCCCATGTCTC 2721
Db 890 rProValThrGluSerCysHisLeuSerSerValAlaPheValProCysLeu 907

RESULT 13

US-10-295-027-484
; Sequence 484, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 484
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-484

Alignment Scores:
Pred. No.: 0 Length: 883
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 68.47% Indels: 0
DB: 15 Gaps: 0
US-10-751-736-21 (1-2724) x US-10-295-027-484 (1-883)
Qy 859 CATTTCATGACAAATCCCAATTTGTTGGGAGATCTGCTTTCAACATTTACCTGAA 918
Db 263 HisPheTyrlAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 282
Qy 919 CTAAAGAACACTGACTCTGAATGTGCTCAAAATAACTGAATTTCTGTATTTAACTGGGA 978
Db 283 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly 302
Qy 979 ACTCAAAACCTGGAGACTCTGACTTTAACTGGAGCAGATCTCATCTCTCCCTCAACACC 1038
Db 303 ThrAlaAsnLeuGluSerLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 322
Qy 1039 GTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCTTACAAACCTATTAGAAGAT 1098
Db 323 ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrlAsnLeuLeuGluAsp 342
Qy 1099 TTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGACCTTAAGACATATGAATC 1158
Db 343 LeuProSerPheSerValCysGlnLysLeuGlnLysIleAspLeuArgHisAsnGluIle 362
Qy 1159 TACGAATTAAGCTTGACACTTCCAGCAGTTGCTTAGCCTCCGATCCGATCGCTGAATTTGCT 1218
Db 363 TyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAla 382
Qy 1219 TGGAAACAAAATTGCTATTATTACCCCAATGCAATTTCCACTTTGGCCATCCCTAATAAG 1278
Db 383 TrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleLys 402
Qy 1279 CTGGACCTATCGTCCAAACCTCTGCTCTTTCTCTATAACTGGGTACATGGTTTAACT 1338
Db 403 LeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThr 422
Qy 1339 CACTTAATTAACAGGAATCATGCTTACAGAGCTTGATATCATCTGAAAACTTTCCA 1398
Db 423 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnLeuPro 442
Qy 1399 GAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGTGTGATTTGGAGTGTGAG 1458
Db 443 GluLeuLysValIleGluMetProTyrlAlaTyrlGlnCysCysAlaPheGlyValCysGlu 462
Qy 1459 AATCCCTATAAGATTTCTAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1518
Db 463 AsnAlaTyrlLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 482
Qy 1519 CATTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGTGAACCTTTGAAGATTTCTG 1578
Db 483 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 502
Qy 1579 CTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGAGTGTGTTCACCTTCCCCAGGC 1638
Db 503 LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 522
Qy 1639 CCCTTCAAAACCTGTGAACACCTGCTGTGCTGGCTGTGATCAGAAATTTGAGTGTGAGACC 1698
Db 523 ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
Qy 1699 ATAGCAGTTCTGGCAGCTTACTTTGTAATGCTTTGTTGAGCTTCAACAGTTTTTCAGATCCCCT 1758
Db 543 IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
Qy 1759 CTGTACATTTCCCCCATTAACCTGTTAATTTGGGCTCATCGCAGCAGTGAACATGCTCAG 1818
Db 563 LeuTyrlSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
Qy 1819 GGAGTCTCCAGTCCGCTGCTGGCTGTGATGCGTTTCACTTTTGGCAGCTTTTGCAGCA 1878
Db 583 GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602

```
QY 1879 CATTGTCCTGCTGGGAGAAATGGGGTTGGTTGCCATGTCATTTGCTTTTGTCCATTTT 1938
Db 603 HisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValileGlyPheLeuSerilePhe 622
QY 1939 GCTTCAGAAATCATCTCTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGTCTCTGNG 1998
Db 623 AlaSerGluSerSerValPheLeuLeuThrLeuAlaLeuGluArgGlyPheSerVal 642
QY 1999 AAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTTG 2058
Db 643 LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValileLeu 662
QY 2059 CTCTGTGCCCTGCTGCTTACCATGGCGCAGTTCCTCTGCTGGTGGCAGCAAGTAT 2118
Db 663 LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlySerLysTyr 682
QY 2119 GCGCCTCCCTCTCTGCTGCTCTTGGGGAGCCAGCAGCAGCATGGCTACATG 2178
Db 683 GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 702
QY 2179 GTGCTCTCATCTTGTCTCAATTCCTTTCCTCATGATGACCATTCCTACACCAAG 2238
Db 703 ValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMetThrileAlaTyrThrLys 722
QY 2239 CTCTACTGCAATTTGCAAGGAGAGACTGGAGATATTTGGAGTCTCTATGTTAA 2298
Db 723 LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnileTrpAspCysSerMetValLys 742
QY 2299 CACATGTCCTGCTCTTACCACCTGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2358
Db 743 HisileAlaLeuLeuLeuPheThrAsnCysileLeuAsnCyProValAlaPheLeuSer 762
QY 2359 TTCTCTCTTTTAATAACCTTACATTTATCAGTCTGAAAGTAATTAAGTTTATCTCTG 2418
Db 763 PheSerSerLeuileAsnLeuThrPheileSerProGluValileLysPheileLeuLeu 782
QY 2419 GTGTAGTCCCATCTCTGCTGCTCAATCCCTCTCTACATCTTGTTCATCTCTAC 2478
Db 783 ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrileLeuPheAsnProHis 802
QY 2479 TTTTAGGAGGACTGGTGGCTGAGAAAGCAAACTAGCTGCTGGCAGATCAAAACAC 2538
Db 803 PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis 822
QY 2539 CCAAGCTGATGTCATTAATCTGATGATGTGCAAAAACAGTCTGCTGACCTCAACTCAA 2598
Db 823 ProSerLeuMetSerileAsnSerAspValGluLysGlnSerCysAspSerThrGln 842
QY 2599 GCCTTGGTAACTTTACAGCTCCAGCATCTATGACCTGCTCCCTCCAGTTCCGTGCA 2658
Db 843 AlaLeuValThrPheThrSerSerSerileThrTyrAspLeuProProSerSerValPro 862
QY 2659 TCACAGCTTATCCAGTGAAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2718
Db 863 SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 882
QY 2719 CTC 2721
Db 883 Leu 883
RESULT 14
US-10-295-027-1330
; Sequence 1330, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezzi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
```

```
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1330
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1330

Alignment Scores:
Pred. No.: 0 Length: 883
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.47% Indels: 0
DB: 15 Gaps: 0

US-10-751-736-21 (1-2724) x US-10-295-027-1330 (1-883)
QY 859 CATTTCATGCAATCCCATCCATTCATTCCTGGAGATCTGCTTTCAACATTTACTCGAA 918
Db 263 HisPheTyrAspAsnProileGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 282
QY 919 CTAAGAACACCTGACTCTGAATGGTCCCTCACAAATAAATGAATTTCTGATTTAACTGGA 978
Db 283 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnileThrGluPheProAspLeuThrGly 302
QY 979 ACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCATCTCTTCTCAACACC 1038
Db 303 ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnileSerSerLeuProGlnThr 322
QY 1039 GTCTGCAATCAGTTACCTAATCTCCAAGTCTGCTAGATCTGCTTACCACTATTAGAGAT 1098
Db 323 ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAsp 342
QY 1099 TTACCAGTCTTTTTCAGTCTGCCAAAGCTTCAGAAAATTCAGCAATATGAACATATGAATC 1158
Db 343 LeuProSerPheSerValCysGlnLysLeuGlnLysileAspLeuArgHisGlnLysile 362
QY 1159 TACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTCCGATCCGCTGAATTTGGCT 1218
Db 363 TyrGluileLysValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAla 382
QY 1219 TGGAAACAAATTCGCTATTATTACCCCAATTCGCTTTCCACTTTGCCATCCCTAATAAG 1278
Db 383 TrpAsnLysileAlaileileHisProAsnAlaPheSerThrLeuProSerLeuileLys 402
```

1279 CTGGACCTATCGTCCAACTCTGTGCTCTTTCTCTATAAATGGGTACATGCTTAACT 1338
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
403 LeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThr 422
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1339 CACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTGAAAACTTTCCA 1398
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
423 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 442
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1399 GAATCAAGGTTATAGAAATGCCCTTATGCTTACCAAGTCTGTGCAATTTGGAGTGTGAG 1458
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
443 GluLeuLysValIleGluMetProfyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 462
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1459 AATGCTATAGATTTCTAATCAATGGAATAAGGTGACACAGCAGATGACGACCTT 1518
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
463 AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 482
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1519 CATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACTTGAAGATTTCTCTG 1578
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
483 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 502
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1579 CTTGACTTTGAGGAGACCTGAAAGCCCTTCAATCAGTGCAGTGTTCACCTTCCCCAGGC 1638
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
503 LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 522
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1639 CCCTTCAACCCCTGTGAACACCTGCTGATGGCTGGCTGATCAGAAATGGAGTGGACC 1698
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
523 PropheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1699 ATAGCAGTCTGGCATTACTTGTAACTGTTGGTGACTTCAACAGTTTTCAGATCCCT 1758
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
543 IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1759 CTGTACATTTCCCCATTTAACTGTTAATGGGTCATCGCAGCAGTGAAACATGCTCAG 1818
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
563 LeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1819 CGAGTCTCCAGTCCGCTGGCTGGTGGTGGATGCTGCTTGGCAGCTTTGGCAGCA 1878
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
583 GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1879 CATGTGCTGTGGGAGAAATGGGCTTGGTGCATGTCATTGTTTGTTCATTTT 1938
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
603 HisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe 622
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1939 GCTTCAGAAATCATCTGTTTCTGTTACTCTGGCAGCCCTGGAGCGTGGGTCTCTGTG 1998
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
623 AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal 642
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1999 AATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTGAAAGTAATCATTTTG 2058
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
643 LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeu 662
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2059 CTCTGTGCCCTGTGCTGGCTTGACCATGGCCGAGTTCCTGCTGGGTGGCAGCAAGPAT 2118
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
663 LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr 682
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2119 GCGCCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGAGCCCGACCATGGCTTACATG 2178
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
683 GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 702
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2179 GTCGCTCTCATCTGCTCAATTCCTTTGCTTCCATGATGACCATTTGCCCTACACCAAG 2238
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
703 ValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys 722
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2239 CTCTACTGCAATTTGGCAAGGAGACCTGAGAAATATTGGGACTGCTCTATGTAATAA 2298
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
723 LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys 742
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2299 CACATTGCCCTGTGCTCTTCAACCAACTGCAATCTTAAACCTGCTGGCTTTCTGTGCC 2358
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
743 HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer 762

Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 2359 TTCTCCTCTTTAATAAACCTTACATTTTATCAGTCTCTGAAGTAATAAGTTTATCTCTCTG 2418
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
763 PheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeu 782
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2419 GTGCTAGTCCACCTTCTCTGATGTCTCAATCCCCTTCTTACATCTTGTTCATCTCCAC 2478
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
783 ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis 802
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2479 TTTAAGAGGATCTGGTGAGCTGAGAAAGCAAACTACGCTGCGACAAGATCAAAACAC 2538
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
803 PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis 822
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2539 CCAAGCTTGTGATGCTCAATTAACCTCTGATGATGCTCAAAAAAGTCTGTGACTCAACTCAA 2598
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
823 ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln 842
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2599 GCCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTGCTCCAGTTCGGTGCCA 2658
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
843 AlaLeuValThrPheThrSerSerSerIleThrTyrAspLeuProProSerSerValPro 862
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2659 TCACAGCTTATCAGTACGACTGAGAGCTGCCATCTTCTCTGTGGCATTTGTCCCATGT 2718
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
863 SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 882
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2719 CTC 2721
Db |||||
883 Leu 883
RESULT 15
US-10-173-999-28
; Sequence 28, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-28
Alignment Scores:
Pred. No.: 0 Length: 883
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.47% Indels: 0
DB: 15 Gaps: 0
US-10-751-736-21 (1-2724) x US-10-173-999-28 (1-883)
Qy 859 CATTCTTATGATCCATCCCATTTTGTGGGAGATCTGCTTTTCAACATTTACTTGAA 918
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
263 HisPheTyrAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 282
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
919 CTAAGAACACTGACTCTGAAATGGTGCCTCACAATAAATGAATTTTCCTGATTTAACTGGA 978

|||||
283 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly 302
QY ACTGCAAACTGAGAGTCTGACTTAACTGGAGCAGACATCTCATCTCTCTCAAAAC 1038
Db ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 322
QY GTCTGCAATCAGTACTTAATCTCAAGTCTAGATCTGTCTTAAACATTAATAGAGAT 1098
Db ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAsp 342
QY TTACCCAGTTTTTCACTGCGCAAAAGCTTCGAAATTCAGCTAGACATAGAAATC 1158
Db LeuProSerPheSerValCysGlnIysLeuGlnIysIleAspLeuArgHisAsnGluIle 362
QY TAGCAAAATTAAGTTCACACTTTCCACAGAGTTGCTTAGCTCCGATCGCTGAATTTGGCT 1218
Db TyrGluIleValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAla 382
QY TGAACAAAATTCGCTATTATTCACCCCAATGCAATTTTCCATTTGCCATCCCTATAAAG 1278
Db TrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleIys 402
QY CTGGACCTATCGTCCAACTCTGCTGCTCTTTCTTATTAACCTGGGTACATGGTTAACT 1338
Db LeuAspLeuSerSerAsnLeuLeuSerPhePheProIleThrGlyLeuHisGlyLeuThr 422
QY CACTTAAATTAACAGAAATCATGCTTACAGAGCTTCATATCATCTGAAACTTTCCA 1398
Db HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 442
QY GAACCAAGGTATAGAAATGCTTATGCTTACAGTCTGCTGCAATTTGGAGTGTGAG 1458
Db GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 462
QY ATGCTTAAAGATTTTCAATCAATGGAATAAGGTGCAACAGCAGTATGACGACCTT 1518
Db AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 482
QY CATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGTGACCTTCAAGATTTCTGT 1578
Db HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuPheLeu 502
QY CTGACTTTGAGAAAGACTGAAAGCCCTTCATTCAGTCGAGTTCACCTTCCCAGGC 1638
Db LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 522
QY CCTTCAAACTGTGTGAACACCTGCTGTGATGCTGCTGCTCATCAGATTGGAGTGGACC 1698
Db ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
QY ATAGCAGTTCTGCACTTACTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1758
Db IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
QY CTGTACATTTCCCACTTAACTGTTAAATTTGGGGTCATCCGACAGTGAACATGCTCAG 1818
Db LeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
QY GGAGTCTCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1878
Db GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602
QY CATGGTCTGTTGGGAGAAATGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1938
Db HisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe 622
QY GCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGCGGTGCTCTGTG 1998
Db AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal 642
QY AAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCTGAAAGTAAATCATTTTG 2058
|||||

Db 643 LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleIleLeu 662
QY CTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2118
Db LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr 682
QY GCGGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2178
Db GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 702
QY GTGCGTCTCATCTGCTCAATTCCTTTCCTCATGATGACCATTCCTGACCAAG 2238
Db ValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys 722
QY CTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTGGGACTGCTCTTATGGTAA 2298
Db LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys 742
QY GACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2358
Db HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer 762
QY TTCTCTCTCTTAAATAAACCTTATACATTTTACGTCCTGAAGTAATTAAGTTTATCTTCT 2418
Db PheSerSerLeuLeuLeuAsnLeuThrPheIleSerProGluValIleLysPheIleLeu 782
QY GTGGTAGTCCCACTTCTGCTGCTGCTCAATCCCTTCTTACATCTTGTTCATCTCCTAC 2478
Db ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis 802
QY TTTAAGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2538
Db PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis 822
QY CCAAGCTTGATGCTCAATTAACCTCATGATGTCGAAAGAACACTGCTGCTGCTCAACTCA 2598
Db ProSerLeuMetSerIleAsnSerAspAspValGluLysGlnSerCysAspSerThrGln 842
QY GCCTTGGTAAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCTGCTGCTGCTGCTGCT 2658
Db AlaLeuValThrPheThrSerSerSerIleThrTyrAspLeuProProSerSerValPro 862
QY TCACAGCTTATCCAGTCACTGAGAGCTGCGCATCTTCTCTGCTGCTGCTGCTGCTGCTGCT 2718
Db SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 882
QY CTC 2721
Db 883 Leu 883

Search completed: July 13, 2005, 04:08:29
Job time : 331.5 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C	1	2724	100.0	2724	6	CQ721366 Sequence
	2	2724	100.0	2724	6	CQ766844 Sequence
	3	2724	100.0	2724	6	AR308745 Sequence
	4	2724	100.0	2724	6	AX549136 Sequence
	5	2724	100.0	2880	6	AX658241 Sequence
	6	2724	100.0	2880	9	AF062006 Homo sapi
	7	2724	100.0	4559	6	BD075815 G protein
	8	2724	100.0	4570	6	AR447625 Sequence
	9	2622	96.3	2724	6	AR308754 Sequence
	10	2571	94.4	2724	9	AF061444 Homo sapi
C	11	1842	67.6	2082	6	BD135244 Novel mam
	12	1090	40.0	145165	9	AC078860 Homo sapi
	13	1019	37.4	4095	9	AK075399 Homo sapi
	14	605	22.2	606	6	E42915 Novel gonad
	15	605	22.2	606	6	AX016185 Sequence
C	16	126	7.8	162004	9	AC030816
	17	112	4.6	816	11	BV208561 GPR49 20
C	18	60	2.2	60	6	CQ550152 Sequence
	19	43	1.6	252477	2	AC136417 Rattus no

[illegible]

Db CCTTCCAACTCAGCGTCTTCACTCTTCACTCTAGACCTCAGTATGAACAACATCAGTCAG 240
QY CTGCTCCCGAATCCCTCCGACAGTCTCCGCTTCTGAGGAGATTACGTTGCGGGAAAC 300
Db CTGCTCCCGAATCCCTCCGACAGTCTCCGCTTCTGAGGAGATTACGTTGCGGGAAAC 300
QY GCTCTGACATACATTTCCCAAGGAGCATCTCACTGGCCCTTTACAGTCTTTAAAGTCTTTATG 360
Db GCTCTGACATACATTTCCCAAGGAGCATCTCACTGGCCCTTTACAGTCTTTAAAGTCTTTATG 360
QY CTGCAGAAATAATCAGCTAAGACACGTAACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db CTGCAGAAATAATCAGCTAAGACACGTAACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
QY CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTCAAGTGGC 480
Db CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTCAAGTGGC 480
QY CTGCAATCCCTGAGGCACTCTGGCTGGATGCAATGCGTTAAACAGAAATCCCGTCCAG 540
Db CTGCAATCCCTGAGGCACTCTGGCTGGATGCAATGCGTTAAACAGAAATCCCGTCCAG 540
QY GCTTTTGAAGTATTCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db GCTTTTGAAGTATTCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
QY ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGGTAGTCTTACATCTCCATTAACAAT 660
Db ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGGTAGTCTTACATCTCCATTAACAAT 660
QY AGAATCCACTCCCTGGGAAAGAAATGCTTTCAGTGGCTCCACAGCTTAGACACTTTAGAT 720
Db AGAATCCACTCCCTGGGAAAGAAATGCTTTCAGTGGCTCCACAGCTTAGACACTTTAGAT 720
QY TTAATAACAATAAATCTTGATGAAATTCCTCCACTGCAATTTAGSACACTCTCCAACTTAA 780
Db TTAATAACAATAAATCTTGATGAAATTCCTCCACTGCAATTTAGSACACTCTCCAACTTAA 780
QY GAACTAGAGATTCATAGCAACAAATATCAGGTGCTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db GAACTAGAGATTCATAGCAACAAATATCAGGTGCTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY CCTCTCTTATTAACAATAATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
Db CCTCTCTTATTAACAATAATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
QY TTTCAAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAACAAATACTGAA 960
Db TTTCAAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAACAAATACTGAA 960
QY TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
QY TCATCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAGTCTAGATCTGCT 1080
Db TCATCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAGTCTAGATCTGCT 1080
QY TACAACTTATTAAGAGATTTTACCAAGTCTGCAAAAGCTTTAGGAGTCTGCT 1140
Db TACAACTTATTAAGAGATTTTACCAAGTCTGCAAAAGCTTTAGGAGTCTGCT 1140
QY CTAAGACATTAATGAATCTAGAAATTAAGTGTGACATTTCCAGCAGTGTGCTAGCCTC 1200
Db CTAAGACATTAATGAATCTAGAAATTAAGTGTGACATTTCCAGCAGTGTGCTAGCCTC 1200
QY CGATCGCTGAATTTGCTTTGGAAACAAATTTGCTATTTTACCCCAATGCAATTTTCCACT 1260
Db CGATCGCTGAATTTGCTTTGGAAACAAATTTGCTATTTTACCCCAATGCAATTTTCCACT 1260
QY TTGCACTCCCTAATAAGCTGCACTATCGTCCAAACCTCTGCTTTTCTTAACT 1320
Db TTGCACTCCCTAATAAGCTGCACTATCGTCCAAACCTCTGCTTTTCTTAACT

Db TTGCACTCCCTAATAAGCTGCACTATCGTCCAAACCTCTGCTGCTTTTCTTAACT 1320
QY GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
QY TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440
Db TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440
QY GCATTTGGAGTGTGAGAAATGCCCTAAGAAATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db GCATTTGGAGTGTGAGAAATGCCCTAAGAAATTTCTAATCAATGGAATAAAGGTGACAA 1500
QY AGCAGTATGAGCACTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db AGCAGTATGAGCACTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
QY GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY TGTTTCACTTTCCAGAGCCCTTCAAAACCTGTGAAACACTGTGTGATGGCTGCTGATC 1680
Db TGTTTCACTTTCCAGAGCCCTTCAAAACCTGTGAAACACTGTGTGATGGCTGCTGATC 1680
QY AGAATTTGGAGTGTGAGCACTAGCAGTCTGAGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
Db AGAATTTGGAGTGTGAGCACTAGCAGTCTGAGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
QY ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACTGTTAAATTTGGGGTTCATCGCA 1800
Db ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACTGTTAAATTTGGGGTTCATCGCA 1800
QY GCAGTGAACATGCTCAAGGAGTCTCCAGTCCGCTGCTGGTGGTGGATGCTTCACT 1860
Db GCAGTGAACATGCTCAAGGAGTCTCCAGTCCGCTGCTGGTGGTGGATGCTTCACT 1860
QY TTTGCGAGCTTTGCAACGACATGCTGGTGGGAGAAATGGGGTGGTGGCAATGTCATT 1920
Db TTTGCGAGCTTTGCAACGACATGCTGGTGGGAGAAATGGGGTGGTGGCAATGTCATT 1920
QY GGTTTTTTGTCCATTTTGTCTTCAAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db GGTTTTTTGTCCATTTTGTCTTCAAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
QY GAGCGTGGTGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2040
Db GAGCGTGGTGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2040
QY CTGAAAGTAAATCAATTTTGTCTCTGTCCTCTGCTGCTTGAACATGCGCGCAGTTCCTCTG 2100
Db CTGAAAGTAAATCAATTTTGTCTCTGTCCTCTGCTGCTTGAACATGCGCGCAGTTCCTCTG 2100
QY CTGGGTGGCAGCAAGTATGGCCCTCTCTGCTGCTTGGCTTTTGGGGAGGCC 2160
Db CTGGGTGGCAGCAAGTATGGCCCTCTCTGCTGCTTGGCTTTTGGGGAGGCC 2160
QY AGCAACATGGGCTACATGGTCTCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db AGCAACATGGGCTACATGGTCTCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
QY ACCATTCCTACCAAGCTCTAGTCAATTTTGACAAGGGAGACCTGGAGAAATTTGG 2280
Db ACCATTCCTACCAAGCTCTAGTCAATTTTGACAAGGGAGACCTGGAGAAATTTGG 2280
QY GACTGCTCTATGTAAGCAATTCGCTTGTGCTTCTTCAACCAACTGCACTCTAAACTGC 2340
Db GACTGCTCTATGTAAGCAATTCGCTTGTGCTTCTTCAACCAACTGCACTCTAAACTGC 2340
QY CCTGTGGCTTTCTGCTTCTCTCTTAAATAAACCCTTACATTTATCAGTCTCTGAAGTA 2400
Db CCTGTGGCTTTCTGCTTCTCTCTTAAATAAACCCTTACATTTATCAGTCTCTGAAGTA 2400

Qy	2401	ATTTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCTGCAATGTCCTCAATGCCCTTCTCTAC	2460
Db	2401	ATTTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCTGCAATGTCCTCAATGCCCTTCTCTAC	2460
Qy	2461	ATCTTGTGTTCAATCCTCACTTTTAAAGGAGGATCTGCTGAGCCCTGAGAAAGCAAACTACGTC	2520
Db	2461	ATCTTGTGTTCAATCCTCACTTTAAGAGGATCTGCTGAGCCCTGAGAAAGCAAACTACGTC	2520
Qy	2521	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGATGATGATCGAAAAACAG	2580
Db	2521	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGATGATGATCGAAAAACAG	2580
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGTTAACTTTACAGACTCCAGCATCACTTATGACCTG	2640
Db	2581	TCCTGTGACTCAACTCAAGCCTTGTTAACTTTACAGACTCCAGCATCACTTATGACCTG	2640
Qy	2641	CCTCCCAAGTTCGGTGCCATCACCAAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTTCCCTCT	2700
Db	2641	CCTCCCAAGTTCGGTGCCATCACCAAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTTCCCTCT	2700
Qy	2701	GTGGCAATTTGTGCCATGTCTCTAA	2724
Db	2701	GTGGCAATTTGTGCCATGTCTCTAA	2724

RESULT 2					
CQ766844					
LOCUS					
	CQ766844	2724 bp	DNA	linear	PAT 07-JUN-2004
DEFINITION	Sequence 3 from Patent WO2004005457.				

ACCESSION	CQ766844
VERSION	2
KEYWORDS	GI:48426391
SOURCE	-
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 Colland,F., Barker,N., Clevers,J.C., Gomez,E., van de Wetering,M.L. and Suile,E.S. THE USE OF SPECIFIED TCF TARGET GENES TO IDENTIFY DRUGS FOR THE TREATMENT OF CANCER, IN PARTICULAR COLORECTAL CANCER, IN WHICH TCF/ctbgr:CATENIN/WNT SIGNALLING PLAYS A CENTRAL ROLE
JOURNAL	Patent: WO 2004005457-A 3 15-JAN-2004; Kylis B.V. (NL)
COMMENT	On Jun 7, 2004 this sequence version replaced gi:44909056.
FEATURES	Location/Qualifiers
source	1. .2724

```

source
1..2724
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db xref="taxon:9606"

```

```

ORIGIN
Query Match 100.0%; Score 2724; DB 6; Length 2724;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/wu_xlsl=Caxon:9606

```

Qy	1	ATGGACACCTCCGGCTCGGTGTGCTCCTGTCTTGCCTGTGCTGCTGACGTGGGAC	60
Db	1	ATGGACACCTCCGGCTCGGTGTGCTCCTGTCTTGTGCTGCTGACGTGGGAC	60
Qy	61	GGGGGACGCTCCGAGGTCGGGTGCTGCTGAGGGCTGCCCAACACTGTCATTC	120
Db	61	GGGGGACGCTCTCCGAGGTCGGGTGCTGCTGAGGGCTGCCCAACACTGTCATTC	120
Qy	121	GAGCCCAACGGCAGGATGTTGCTCAGGGTGAGTGTCCGACCTGGGGCTCTCGGAGCTG	180
Db	121	GAGCCCAACGGCAGGATGTTGCTCAGGGTGAGTGTCCGACCTGGGGCTCTCGGAGCTG	180
Qy	181	CTTTCCAACTCAGGGTCTTCACTCTCACTAGACCTCAGTAGTAAACAATCAGTCAG	240
Db	181	CTTTCCAACTCAGGGTCTTCACTCTCACTAGACCTCAGTAGTAAACAATCAGTCAG	240

Qy	241	CTGCTCCGGAATCCCTCGCCAGTCTCGGTCTCTCGAGGAGTTCAGTCTTGGCGGAAC	300
Db	241	CTGCTCCGGAATCCCTCGCCAGTCTCGGTCTCTCGAGGAGTTCAGTCTTGGCGGAAC	300
Qy	301	GCTCTGACATACATTCCTCAAGGAGCAATTCACATGGCCCTTTACAGTCTTAAAGTTCCTTATG	360
Db	301	GCTCTGACATACATTCCTCAAGGAGCAATTCACATGGCCCTTTACAGTCTTAAAGTTCCTTATG	360
Qy	361	CTGCAGAAATATCAGCTTAAGACACGTACCCACAGAACTCTGCGAAGCTTTCGGAAGCCTT	420
Db	361	CTGCAGAAATATCAGCTTAAGACACGTACCCACAGAACTCTGCGAAGCTTTCGGAAGCCTT	420
Qy	421	CAATCCCTGGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Db	421	CAATCCCTGGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Qy	481	CTGCATTCCTCGAGGCACTGTGCTGGATGACAATGCGTTTAAAGAAATCCCGCTCAG	540
Db	481	CTGCATTCCTCGAGGCACTGTGCTGGATGACAATGCGTTTAAAGAAATCCCGCTCAG	540
Qy	541	GCCTTTAGAAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAAACAAAATACACCAC	600
Db	541	GCCTTTAGAAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAAACAAAATACACCAC	600
Qy	601	ATACCGAGCTATGCTTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	660
Db	601	ATACCGAGCTATGCTTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	660
Qy	661	AGAACTCCCTCGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720
Db	661	AGAACTCCCTCGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720
Qy	721	TTAAATTAACAATAACCTTGAATGAATTCCTCCACTGCAATTAGGACACTCTCCAACTTTAAA	780
Db	721	TTAAATTAACAATAACCTTGAATGAATTCCTCCACTGCAATTAGGACACTCTCCAACTTTAAA	780
Qy	781	GAACTAGGATTTCAATAGCAACAATATCAGGTGCAATCTGAGAAAGCAATTTGTAGGCAAC	840
Db	781	GAACTAGGATTTCAATAGCAACAATATCAGGTGCAATCTGAGAAAGCAATTTGTAGGCAAC	840
Qy	841	CTCTCTTTATTAACAATATTTCTATGCAATCCCATCCAAATTTGTTGGGAGATCTGCT	900
Db	841	CTCTCTTTATTAACAATATTTCTATGCAATCCCATCCAAATTTGTTGGGAGATCTGCT	900
Qy	901	TTTCAACATTTACTGAACTAAGAACTGACTCTGAAATGGTGCTCACAATAAAGCTGAA	960
Db	901	TTTCAACATTTACTGAACTAAGAACTGACTCTGAAATGGTGCTCACAATAAAGCTGAA	960
Qy	961	TTTCTGATTTAACTGGAACCTGGAAGTCTGACTTTAACTGGAGCACAGATC	1020
Db	961	TTTCTGATTTAACTGGAACCTGGAAGTCTGACTTTAACTGGAGCACAGATC	1020
Qy	1021	TCATCTCTTCTCAACCGTCTGCAATCAGTTACCTAACTCTCAAGTGCTTAGATCTGCT	1080
Db	1021	TCATCTCTTCTCAACCGTCTGCAATCAGTTACCTAACTCTCAAGTGCTTAGATCTGCT	1080
Qy	1081	TACAACCTATTAGAAATTTACCAGGTTTTTACGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140
Db	1081	TACAACCTATTAGAAATTTACCAGGTTTTTACGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTTCAGGCTC	1200
Db	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTTCAGGCTC	1200
Qy	1201	CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCACCCCAATGCAATTTCCACT	1260
Db	1201	CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCACCCCAATGCAATTTCCACT	1260
Qy	1261	TTGCCATCCCTAATAAAGCTGGACCTATGCTGCCAACTCTCTGCTCTTTTCCTATAACT	1320
Db	1261	TTGCCATCCCTAATAAAGCTGGACCTATGCTGCCAACTCTCTGCTCTTTTCCTATAACT	1320
Qy	1321	GGGTACATGGTTTAACTCACTTTAAATTTAAACAGGAAATCATGCCCTTACAGAGCTTGATA	1380

[illegible]

Db	1441	GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAAATCAAATGGAAATAAAGGTGCAAC	1500
Qy	1501	AGCAGTATGACGACCTTTCAATAAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1501	AGCAGTATGACGACCTTTCAATAAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1560
Qy	1561	GACCTTGAAGATTTCTGTCTGTGACTTTGAGGAGAACTCTGAAAGCCCTTCATTCAAGTGCAG	1620
Db	1561	GACCTTGAAGATTTCTGTCTGTGACTTTGAGGAGAACTCTGAAAGCCCTTCATTCAAGTGCAG	1620
Qy	1621	TGTTCACTTCCCGAGGCCCTTTCAACCCCTGTGAACCTGTGATGGCTGGCTGGCTGATC	1680
Db	1621	TGTTCACTTCCCGAGGCCCTTTCAACCCCTGTGAACCTGTGATGGCTGGCTGGCTGATC	1680
Qy	1681	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1681	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Qy	1741	ACAGTTTTTCAGATCCCTCTCTGTHACATTTCCCCCATTTAAACTGTATAATGCGGTCAFCGCA	1800
Db	1741	ACAGTTTTTCAGATCCCTCTCTGTHACATTTCCCCCATTTAAACTGTATAATGCGGTCAFCGCA	1800
Qy	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACT	1860
Db	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACT	1860
Qy	1861	TTTGGCAGCTTTGCACACATGCTGCTGTGTGGAGAAATGGGGTGTGTTGCCATGTCATT	1920
Db	1861	TTTGGCAGCTTTGCACACATGCTGCTGTGTGGAGAAATGGGGTGTGTTGCCATGTCATT	1920
Qy	1921	GGTTTTTTGFCATTTTTTGTCTCAGAAATCATCTGTTTTCTTGCTTACTCTGGCAGCCCTG	1980
Db	1921	GGTTTTTTGFCATTTTTTGTCTCAGAAATCATCTGTTTTCTTGCTTACTCTGGCAGCCCTG	1980
Qy	1981	GACGTGGGTTCTCTGTGAAATATCTCGAAAATTTTGAAAACGAAAGCTCATTTCCTAGC	2040
Db	1981	GACGTGGGTTCTCTGTGAAATATCTCGAAAATTTTGAAAACGAAAGCTCATTTCCTAGC	2040
Qy	2041	CTGAAAGTAAATCATTTTTGTCTGTGCTGCTGGCTTGACCAATGCGCGCAGTTCCCTCG	2100
Db	2041	CTGAAAGTAAATCATTTTTGTCTGTGCTGCTGGCTTGACCAATGCGCGCAGTTCCCTCG	2100
Qy	2101	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTCTGCTGCTTGTGCTTTTGGGGAGCCCC	2160
Db	2101	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTCTGCTGCTTGTGCTTTTGGGGAGCCCC	2160
Qy	2161	AGCACCATGGCTACATGGTTCGCTCATCTTGCTCAATTCCTTTGCTTCTCTCATGATG	2220
Db	2161	AGCACCATGGCTACATGGTTCGCTCATCTTGCTCAATTCCTTTGCTTCTCTCATGATG	2220
Qy	2221	ACCAATTGCTTACACCAAGCTCTACTGCAATTTTGGACAGGAGACCTCGAGAAATATTGG	2280
Db	2221	ACCAATTGCTTACACCAAGCTCTACTGCAATTTTGGACAGGAGACCTCGAGAAATATTGG	2280
Qy	2281	GACTGCTCTATGGTAAAAACATATGCCCTGTGCTCTTCCACCACTGCATCTCTAAACTGC	2340
Db	2281	GACTGCTCTATGGTAAAAACATATGCCCTGTGCTCTTCCACCACTGCATCTCTAAACTGC	2340
Qy	2341	CTGTGGCTTTCTTGTCCTTCTCTTTAAATAAACCTTACATTTATCATGCTCTGAAGTA	2400
Db	2341	CTGTGGCTTTCTTGTCCTTCTCTTTAAATAAACCTTACATTTATCATGCTCTGAAGTA	2400
Qy	2401	ATTAAAGTTTATCCTTCTGTGGTGTAGTCCCACTTCTGCAATGCTCAATCCCCTTCTCTAC	2460
Db	2401	ATTAAAGTTTATCCTTCTGTGGTGTAGTCCCACTTCTGCAATGCTCAATCCCCTTCTCTAC	2460
Qy	2461	ATCTTGTTCATCTCACTTTTAAAGGAGGATCTGGTGAGCCTCGAGAAAGCAAACTACGTC	2520
Db	2461	ATCTTGTTCATCTCACTTTTAAAGGAGGATCTGGTGAGCCTCGAGAAAGCAAACTACGTC	2520
Qy	2521	TGACAAAGATCAAAAACCCCAAGCTTGATGTCAAATTAATCTGTATGATGTGAAAACAG	2580
Db	2521	TGACAAAGATCAAAAACCCCAAGCTTGATGTCAAATTAATCTGTATGATGTGAAAACAG	2580


```
QY 1561 GACCTTGAAGATTCTGCTGTGATCTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGTCAG 1620
DB 1561 GACCTTGAAGATTCTGCTGTGATCTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGTCAG 1620
QY 1621 TGTTCACCTTTCCCGAGGCCCTTTCAAAACCTGTGAAACCTGCTGTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACCTTTCCCGAGGCCCTTTCAAAACCTGTGAAACCTGCTGTGATGGCTGGCTGATC 1680
QY 1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGACATCTTACTTGTAAATGCTTTGGTGAATCA 1740
DB 1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGACATCTTACTTGTAAATGCTTTGGTGAATCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAATTTGGGGTTCATCGCA 1800
DB 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAATTTGGGGTTCATCGCA 1800
QY 1801 GCAATGAACATGCTCAGGGAGTCTCCAGTGGCGTCTGCTGCTGTGATGGATCGTTCACT 1860
DB 1801 GCAATGAACATGCTCAGGGAGTCTCCAGTGGCGTCTGCTGCTGTGATGGATCGTTCACT 1860
QY 1861 TTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
DB 1861 TTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
QY 1921 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGTCAGCCCTG 1980
DB 1921 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGTCAGCCCTG 1980
QY 1981 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAAACCAAGCTCCATTTTCTAGC 2040
DB 1981 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAAACCAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACATGGGTACATGCTGCTCTCATCTGCTCTCATCTGCTCTCATCTGCTCTCATCTG 2220
DB 2161 AGCACATGGGTACATGCTGCTCTCATCTGCTCTCATCTGCTCTCATCTGCTCTCATCTG 2220
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATTTGG 2280
DB 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATTTGG 2280
QY 2281 GACTGCTCTATGGTAAACACATTTGCTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
DB 2281 GACTGCTCTATGGTAAACACATTTGCTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
QY 2341 CCGTGGCTTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 2400
DB 2341 CCGTGGCTTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 2400
QY 2401 ATTTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
DB 2401 ATTTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 ATCTGTTCAATCTCTCACTTTAAGGAGGATCTGGTGGCTTGGAGCTTGGAGCTTGGAGCT 2520
DB 2461 ATCTGTTCAATCTCTCACTTTAAGGAGGATCTGGTGGCTTGGAGCTTGGAGCTTGGAGCT 2520
QY 2521 TGGACAGATCAAAACCAAGCTTGTATGTAATTAATCTGATGATGATGATGATGATGATGATG 2580
DB 2521 TGGACAGATCAAAACCAAGCTTGTATGTAATTAATCTGATGATGATGATGATGATGATGATG 2580
QY 2581 TCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 2581 TCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
```

```
QY 2641 CTTCCAGTTCGTCGCCATCACCAGCTTATCCAGTGAAGAGCTGCCATCTTTCCTCT 2700
DB 2641 CTTCCAGTTCGTCGCCATCACCAGCTTATCCAGTGAAGAGCTGCCATCTTTCCTCT 2700
QY 2701 GTGGCATTTGTCCCATGCTCTAA 2724
DB 2701 GTGGCATTTGTCCCATGCTCTAA 2724

RESULT 5
AX658241
LOCUS AX658241
DEFINITION Sequence 157 from Patent WO0300928.
ACCESSION AX658241
VERSION AX658241.1 GI:29160776
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
        Petersen,M.W. and Elsner,H.I.
TITLE Methods for identification of cancer cell surface molecules and
        cancer specific promoters, and therapeutic uses thereof
JOURNAL Patent: WO 0300928-A 157 03-JAN-2003;
        Odin Medical A/S (DK)
FEATURES             Location/Qualifiers
     source            1..2880
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 2724; DB 6; Length 2880;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACACTCCCGGCTCGGTGTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
DB 49 ATGGACACTCCCGGCTCGGTGTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108
QY 61 GGGGGCAGCTCTCCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCAATGC 120
DB 109 GGGGGCAGCTCTCCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCAATGC 168
QY 121 GAGCCGAGCGCAGATGTTGCTCAGGGTGACCTGCTCCGACCTGGGCTCTCGGAGCTG 180
DB 169 GAGCCGAGCGCAGATGTTGCTCAGGGTGACCTGCTCCGACCTGGGCTCTCGGAGCTG 228
QY 181 CTTTCCAACTCAGCGTCTTCACTCTTACCTTAGACCTCAGTATGAAACAAATCAGTCAG 240
DB 229 CTTTCCAACTCAGCGTCTTCACTCTTACCTTAGACCTCAGTATGAAACAAATCAGTCAG 288
QY 241 CTGCTCCCGAATCCCTGCTCCAGTCTCCGCTTCTGAGAGGATTAAGTCTTGGCGGAAAC 300
DB 289 CTGCTCCCGAATCCCTGCTCCAGTCTCCGCTTCTGAGAGGATTAAGTCTTGGCGGAAAC 348
QY 301 GCTCTGACATACATTTCCCAAGGGAGCTTCACTGGCCCTTACAGTCTTAAAGTCTTATG 360
DB 349 GCTCTGACATACATTTCCCAAGGGAGCTTCACTGGCCCTTACAGTCTTAAAGTCTTATG 408
QY 361 CTGCAGATAATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
DB 409 CTGCAGATAATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 468
QY 421 CAATCCCTGCTCTGGATGCTAAACATCAGCTATGTCGCCCCCAAGCTGTTTCAAGTGC 480
DB 469 CAATCCCTGCTCTGGATGCTAAACATCAGCTATGTCGCCCCCAAGCTGTTTCAAGTGC 528
QY 481 CTGCATTTCCCTGAGGACCTGTGGCTGATGATGATGATGATGATGATGATGATGATGATG 540
DB 529 CTGCATTTCCCTGAGGACCTGTGGCTGATGATGATGATGATGATGATGATGATGATGATG 588
```


QY	541	GCCTTTAGAGTTTATCGGCATTGCAAGCATGACCTTGGCCCTGAACAAATATACCCAC	600
Db	589	GCCTTTAGAGTTTATCGGCATTGCAAGCATGACCTTGGCCCTGAACAAATATACCCAC	648
QY	601	ATACCAGACTATGCCTTTGGAAAACTCTCCAGCTTGGTAGTCTTACACTCTCCATAACAAT	660
Db	649	ATACCAGACTATGCCTTTGGAAAACTCTCCAGCTTGGTAGTCTTACACTCTCCATAACAAT	708
QY	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGAT	720
Db	709	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGAT	768
QY	721	TTAAATTAACAATAACCTTGGTAATTTCCCACTGCAATTTAGACACTCTCCAACTTAAA	780
Db	769	TTAAATTAACAATAACCTTGGTAATTTCCCACTGCAATTTAGACACTCTCCAACTTAAA	828
QY	781	GAACCTAGGATTTATAGCAAAATATCAGGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db	829	GAACCTAGGATTTATAGCAAAATATCAGGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC	888
QY	841	CCCTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	900
Db	889	CCCTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	948
QY	901	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGCTGCTCACAATTAACCTGAA	960
Db	949	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGCTGCTCACAATTAACCTGAA	1008
QY	961	TTTTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db	1009	TTTTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1068
QY	1021	TCATCTCTTCCCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAGTCTAGATCTGCT	1080
Db	1069	TCATCTCTTCCCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAGTCTAGATCTGCT	1128
QY	1081	TACAACTTATAGAGATTTACCGAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC	1140
Db	1129	TACAACTTATAGAGATTTACCGAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC	1188
QY	1141	CTAAGACATTAATGAAATCTACGAAATTAAGTTTGAACCTTTCCAGCAGTCTGCTAGCCTC	1200
Db	1189	CTAAGACATTAATGAAATCTACGAAATTAAGTTTGAACCTTTCCAGCAGTCTGCTAGCCTC	1248
QY	1201	CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT	1260
Db	1249	CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT	1308
QY	1261	TTGCCATCCCTTAATAAGCTGGACCTATGCTCCAACTCTGCTGCTTTTCCCTATTAAC	1320
Db	1309	TTGCCATCCCTTAATAAGCTGGACCTATGCTCCAACTCTGCTGCTTTTCCCTATTAAC	1368
QY	1321	GGGTTTACATGGTTTAACTCACCTTAAATTTAAGGAAATCATGCTTACAGAGCTTTGATA	1380
Db	1369	GGGTTTACATGGTTTAACTCACCTTAAATTTAAGGAAATCATGCTTACAGAGCTTTGATA	1428
QY	1381	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTTCAAGTGTGT	1440
Db	1429	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTTCAAGTGTGT	1488
QY	1441	GCATTTGGAGTGTGTGAGATGCCCTTAAGATTTCTTAATCAATGGAATTAAGGTTGACAAAC	1500
Db	1489	GCATTTGGAGTGTGTGAGATGCCCTTAAGATTTCTTAATCAATGGAATTAAGGTTGACAAAC	1548
QY	1501	AGCAGTATGGAGCACTTTCATAGAAAGATGTGGAATGTTTTCAGGCTCAAGATGAAAGT	1560
Db	1549	AGCAGTATGGAGCACTTTCATAGAAAGATGTGGAATGTTTTCAGGCTCAAGATGAAAGT	1608
QY	1561	GACCTTGAGATTTCTGCTTGATCTTTTGAGGAAGACCTTGAAAGCCCTTCAATTCAGTGAG	1620
Db	1609	GACCTTGAGATTTCTGCTTGATCTTTTGAGGAAGACCTTGAAAGCCCTTCAATTCAGTGAG	1668

QY	1621	TGTTTACCTTCCAGGCCCTTTCAAACTCTGTGAACACCTGTCTGTATGGCTGGCTGATC	1680
Db	1669	TGTTTACCTTCCAGGCCCTTTCAAACTCTGTGAACACCTGTCTGTATGGCTGGCTGATC	1728
QY	1681	AGAAATCGAGTGTGGACCATAGCAGATTCTGGCACTTACTTGTGTAAATGCTTTGGTGAATTCA	1740
Db	1729	AGAAATCGAGTGTGGACCATAGCAGATTCTGGCACTTACTTGTGTAAATGCTTTGGTGAATTCA	1788
QY	1741	ACAGTTTTAGATCCCTCTGTGTACATTTTCCCAATTAACCTGTTAAATTTGGGGTCAATCGCA	1800
Db	1789	ACAGTTTTAGATCCCTCTGTGTACATTTTCCCAATTAACCTGTTAAATTTGGGGTCAATCGCA	1848
QY	1801	GCAGTGAACATGCTTCAGGGAGTCTCCAGTGCCTGTGCTGGTGTGGAATGCGTTTCACT	1860
Db	1849	GCAGTGAACATGCTTCAGGGAGTCTCCAGTGCCTGTGCTGGTGTGGAATGCGTTTCACT	1908
QY	1861	TTTGGCAGCTTTGCAACGACATGCTGCTGGGAGAAATGGGGTGGTGGCCATGTCATT	1920
Db	1909	TTTGGCAGCTTTGCAACGACATGCTGCTGGGAGAAATGGGGTGGTGGCCATGTCATT	1968
QY	1921	GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTCTGGCAGCCCTG	1980
Db	1969	GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTCTGGCAGCCCTG	2028
QY	1981	GAGGCTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040
Db	2029	GAGGCTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2088
QY	2041	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGTGSCCTTGACCATGGCGGAGTTTCCCTG	2100
Db	2089	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGTGSCCTTGACCATGGCGGAGTTTCCCTG	2148
QY	2101	CTGGGTGGCAGCAGTATGGGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC	2160
Db	2149	CTGGGTGGCAGCAGTATGGGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC	2208
QY	2161	AGCAACCATGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATG	2220
Db	2209	AGCAACCATGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATG	2268
QY	2221	ACCAATGCTCTACCAAGCTCTACTGCAATTTTGAAGGGAGACCTGGAGAAATTTTGG	2280
Db	2269	ACCAATGCTCTACCAAGCTCTACTGCAATTTTGAAGGGAGACCTGGAGAAATTTTGG	2328
QY	2281	GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCACTGCTCATCTTAAACTGC	2340
Db	2329	GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCACTGCTCATCTTAAACTGC	2388
QY	2341	CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAAGTA	2400
Db	2389	CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAAGTA	2448
QY	2401	ATTAAGTTTATCTCTTGTGGTGTAGTCCCACTTCTGCTGCTCAATCTCAATCTCTCTAC	2460
Db	2449	ATTAAGTTTATCTCTTGTGGTGTAGTCCCACTTCTGCTGCTCAATCTCAATCTCTCTAC	2508
QY	2461	ATCTGTTTCAATCTCTCACTTTTAAAGGAGATCTGGTGGAGCCTGAGAAAGCAAACTACGTC	2520
Db	2509	ATCTGTTTCAATCTCTCACTTTTAAAGGAGATCTGGTGGAGCCTGAGAAAGCAAACTACGTC	2568
QY	2521	TGGCAAGATCAAAACACCCCAAGCTTGTATGTCAAATTAATCTGTATGTCTGAAAAACAG	2580
Db	2569	TGGCAAGATCAAAACACCCCAAGCTTGTATGTCAAATTAATCTGTATGTCTGAAAAACAG	2628
QY	2581	TCCTGTGACTCAACTCAAGCCTTGTGTAACTTTTACAGCTCCAGCATCATTATGACCTG	2640
Db	2629	TCCTGTGACTCAACTCAAGCCTTGTGTAACTTTTACAGCTCCAGCATCATTATGACCTG	2688
QY	2641	CCTCCAGTTCGGTGGCCATCAACAGCTTATCCAGTGTAGTGTGCTGCTTCTCTCT	2700
Db	2689	CCTCCAGTTCGGTGGCCATCAACAGCTTATCCAGTGTAGTGTGCTGCTTCTCTCT	2748
QY	2701	GTGGCAATTTGTCCCATGTCTCTAA	2724

Qy	1361	TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCCTCGTCGTCCTTTTCCCTAAACT	1320
Db	1309	TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCCTCGTCGTCCTTTTCCCTAATACT	1368
Qy	1321	GGGTTACATGGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTACAGAGCTTGATA	1380
Db	1369	GGGTTACATGGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTACAGAGCTTGATA	1428
Qy	1381	TCATCTGAAACCTTTCAGAGACTCAAGGTTATAGAATGCTTATGCTTACACAGTCTGT	1440
Db	1429	TCATCTGAAACCTTTCAGAGACTCAAGGTTATAGAATGCTTATGCTTACACAGTCTGT	1488
Qy	1441	GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTCTAATCAATGGGAATAAAGGTGACAAC	1500
Db	1489	GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTCTAATCAATGGGAATAAAGGTGACAAC	1548
Qy	1501	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGGAAGCCCTTCATTCAGTGCAG	1620
Db	1609	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGGAAGCCCTTCATTCAGTGCAG	1668
Qy	1621	TGTTTCACTTCCCAAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC	1680
Db	1669	TGTTTCACTTCCCAAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC	1728
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCATTACTTTGTAATGCTTTGGTGACTTCA	1740
Db	1729	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCATTACTTTGTAATGCTTTGGTGACTTCA	1788
Qy	1741	ACAGTTTTAGATCCCTCTGTACATTTCCGCCATTAACCTGTAAATGGGGTCATGCGCA	1800
Db	1789	ACAGTTTTAGATCCCTCTGTACATTTCCGCCATTAACCTGTAAATGGGGTCATGCGCA	1848
Qy	1801	GCAGTGAACATGTCACGGGAGTCTCCAGTGCCTGCTGCTGCTGTGGATGCGTTCACT	1860
Db	1849	GCAGTGAACATGTCACGGGAGTCTCCAGTGCCTGCTGCTGCTGTGGATGCGTTCACT	1908
Qy	1861	TTTGGCAGCTTTGACGACATGTTGCTGCTGGTGGAGAAATGGGTGGTTCGCCATGCAAT	1920
Db	1909	TTTGGCAGCTTTGACGACATGTTGCTGCTGGTGGAGAAATGGGTGGTTCGCCATGCAAT	1968
Qy	1921	GGTTTTTGTCCATTTTGTCTTACAGATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	1980
Db	1969	GGTTTTTGTCCATTTTGTCTTACAGATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	2028
Qy	1981	GAGCGTGGTTCTCTGTGAATAATTTCTGCAAAATTTGCAAAAGCTCCATTTTCTAGC	2040
Db	2029	GAGCGTGGTTCTCTGTGAATAATTTCTGCAAAATTTGCAAAAGCTCCATTTTCTAGC	2088
Qy	2041	CTGAAAGTAATCATTTTGTCTGTGCCCTGTGCGCTTGACCAATGGCCGAGTTCCCTCG	2100
Db	2089	CTGAAAGTAATCATTTTGTCTGTGCCCTGTGCGCTTGACCAATGGCCGAGTTCCCTCG	2148
Qy	2101	CTGGGTGGCAGAAATATGGCGCCTCCCTCTCTGCGCTTGGCCTTTTGGGGAGCCCC	2160
Db	2149	CTGGGTGGCAGAAATATGGCGCCTCCCTCTCTGCGCTTGGCCTTTTGGGGAGCCCC	2208
Qy	2161	AGCACCATGGGCTACATGGTCGCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATGATG	2220
Db	2209	AGCACCATGGGCTACATGGTCGCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATGATG	2268
Qy	2221	ACCATTGCCTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTTGGAGAAATTTTGG	2280
Db	2269	ACCATTGCCTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTTGGAGAAATTTTGG	2328
Qy	2281	GACTGCTCTATGGTAAAAACAATTTGCCCTGTGTCTTCTTCAACAACTGCATCTCAAATGCG	2340
Db	2329	GACTGCTCTATGGTAAAAACAATTTGCCCTGTGTCTTCTTCAACAACTGCATCTCAAATGCG	2388

	QY	2341	CCTGTGGCTTTTCTTGTCCTTCTCCTCTTTTAATAAAACCTTACAATTATCACTGCTCGTAGTA	2400
	Dd	2389	CCTGTGGCTTTTCTTGTCCTTCTCCTCTTTTAATAAAACCTTACAATTATCACTGCTCGTAGTA	2448
	QY	2401	ATTAAAGTTTTATCCTTCTTGCTGGGTAGTCCCACTTCCTGCAGTGCTCAATCCCCTTCTCTAC	2460
	Dd	2449	ATTAAAGTTTTATCCTTCTTGCTGGGTAGTCCCACTTCCTGCAGTGCTCAATCCCCTTCTCTAC	2508
	QY	2461	ATCTTGTTCATTCCTCACTTTTAAAGGAGACTCTGGTGAGCCCTGAGAAAAGCAAACCTACGTC	2520
	Dd	2509	ATCTTGTTCATTCCTCACTTTAAGGAGACTCTGGTGAGCCCTGAGAAAAGCAAACCTACGTC	2568
	QY	2521	TGGACAAGATCAAAACACCAGAAGCTTGATGTCAATTAACCTCTGATGATGTGAAAAAACAG	2580
	Dd	2569	TGGACAAGATCAAAACACCAGAAGCTTGATGTCAATTAACCTCTGATGATGTGAAAAAACAG	2628
	QY	2581	TCCTGTGAAGTCAACTCAAGCCTTGTTAAAGCTTTTACCAGCTCCAGCATCACCTATGACCTG	2640
	Dd	2629	TCCTGTGAAGTCAACTCAAGCCTTGTTAAAGCTTTTACCAGCTCCAGCATCACCTATGACCTG	2688
	QY	2641	CCTCCAGTTCCTGCCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTTCTCTCT	2700
	Dd	2689	CCTCCAGTTCCTGCCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTTCTCTCT	2748
	QY	2701	GTCGCAATTTGTCCTCAATGCTCTAA	2724
	Dd	2749	GTCGCAATTTGTCCTCAATGCTCTAA	2772
RESULT 7				
BD075815				
LOCUS		BD075815	4559 bp DNA linear PAT 27-AUG-2002	
DEFINITION		G protein-coupled glycoprotein hormone receptor HG38.		
ACCESSION		BD075815		
VERSION		BD075815.1 GI:22621418		
KEYWORDS		JP 2001517441-A/1.		
SOURCE		unidentified		
ORGANISM		unclassified.		
REFERENCE		1. (bases 1 to 4559)		
AUTHORS		Liu,Q., Berly,W.J. and Macdonald,T.P.		
TITLE		G protein-coupled glycoprotein hormone receptor HG38		
JOURNAL		Patent: JP 2001517441-A 1 OCT-2001;		
		MERCK & CO INC		
COMMENT		OS Unidentified		
		PN JP 2001517441-A/1		
		PD 09-OCT-2001		
		PF 24-SEP-1998 JP 2000512952		
		PR 24-SEP-1997 US 60/059863		
		P1 QINGYUN LIU, WENDY J BERLY, TERENCE P MACDONALD PC		
		C12N15/09,A01K67/027,C07K14/72,C07K16/28,C12N1/15,C12N1/19,PC		
		C12N1/21		
		PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,PC		
		C12N15/00,		
		PC C12N5/00		
		CC Strandedness: Double;		
		CC Topology: Linear;		
		CC G protein-coupled glycoprotein hormone receptor HG38 FH Key		
		Location/Qualifiers		
FT source		1..4559		
FEATURES		/organism='Unidentified'.		
source		Location/Qualifiers		
		1..4559		
		/organism="unidentified"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:32644"		
ORIGIN				
Query Match		100.0%; Score 2724; DB 6; Length 4559;		
Best Local Similarity		100.0%; Pred. No. 0;		
Matches 2724; Conservative		0; Mismatches 0; Indels 0; Gaps 0;		
QY		1 ATGGACACTCCCGGCTGGGTGCTCTCTGTCTTGGCTGTGCTGCTGCAGCTGGCGACC	60	

QY 301 GCTCTGACATACATTTCCCAAGGAGCAATTCAGTGGCCCTTTACAGTCTTTAAAGTTCTTTATG 360
DB |||||
DB 301 GCTCTGACATACATTTCCCAAGGAGCAATTCAGTGGCCCTTTACAGTCTTTAAAGTTCTTTATG 360
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCCTT 420
DB |||||
DB 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCCTT 420
QY 421 CAATCCCTGCGTCTGGATGCTTAACCAATCAGCTATGTCGCCGCCCAAGCTGTTTCAGTGGC 480
DB |||||
DB 421 CAATCCCTGCGTCTGGATGCTTAACCAATCAGCTATGTCGCCGCCCAAGCTGTTTCAGTGGC 480
QY 481 CTGCATTCCTGAGGACCTGTCGCTGATGAGCAATGCTTTAACAAGAAATCCCGTCCAG 540
DB |||||
DB 481 CTGCATTCCTGAGGACCTGTCGCTGATGAGCAATGCTTTAACAAGAAATCCCGTCCAG 540
QY 541 GCTTTTGAAGTTTATTCGGCAATGCAAGCCATGACCTTGGCCCTGAAACAAATAACACCAC 600
DB |||||
DB 541 GCTTTTGAAGTTTATTCGGCAATGCAAGCCATGACCTTGGCCCTGAAACAAATAACACCAC 600
QY 601 ATACAGAGCTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTCTACATCTCCATTAACAAT 660
DB |||||
DB 601 ATACAGAGCTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTCTACATCTCCATTAACAAT 660
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
DB |||||
DB 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
QY 721 TTAATAATACAATAAATCTTGGATGAATTCGCCATGCAATTAGGACACTCTCCAAACCTTAAA 780
DB |||||
DB 721 TTAATAATACAATAAATCTTGGATGAATTCGCCATGCAATTAGGACACTCTCCAAACCTTAAA 780
QY 781 GAACTAGAGTTTATAGCAAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB |||||
DB 781 GAACTAGAGTTTATAGCAAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 841 CCTCTCTTATACAAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
DB |||||
DB 841 CCTCTCTTATACAAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
QY 901 TTTCAACATTTTACCTGAATAGAAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 960
DB |||||
DB 901 TTTCAACATTTTACCTGAATAGAAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 960
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
DB |||||
DB 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGATTAACCTAATCTCAAGTGTAGATCTGCT 1080
DB |||||
DB 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGATTAACCTAATCTCAAGTGTAGATCTGCT 1080
QY 1081 TACAACTTATAGAAATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTCAC 1140
DB |||||
DB 1081 TACAACTTATAGAAATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTCAC 1140
QY 1141 CTAAGACATTAATAGAAATCTACGAAATTAAGTTGACATTTCCAGCAGTGTGCTAGCCCTC 1200
DB |||||
DB 1141 CTAAGACATTAATAGAAATCTACGAAATTAAGTTGACATTTCCAGCAGTGTGCTAGCCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
DB |||||
DB 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAGCTGGAACCTATCGTCCAAACCTCTGTCGCTTTTTCCTATAACT 1320
DB |||||
DB 1261 TTGCCATCCCTAATAAGCTGGAACCTATCGTCCAAACCTCTGTCGCTTTTTCCTATAACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCCCTTACAGAGCTTGATA 1380
DB |||||
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCCCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAACCTTTTCCGAACTCAAGGTTATAGAAATGSCCTTATGCTTACCAGTGTGT 1440

DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTATAGAAATGSCCTTATGCTTACCAGTGTGT 1440
QY 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAATTAAGTTTCTAATCAATGAAATAAAGGTGACAAAC 1500
DB |||||
DB 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAATTAAGTTTCTAATCAATGAAATAAAGGTGACAAAC 1500
QY 1501 AGCAGTATGAGACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGGCTCAAGATGAACGT 1560
DB |||||
DB 1501 AGCAGTATGAGACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGGCTCAAGATGAACGT 1560
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAACCTGAAAGCCCTTCATTTCAGTGCAG 1620
DB |||||
DB 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAACCTGAAAGCCCTTCATTTCAGTGCAG 1620
QY 1621 TGTTCACCTTCCAGGCGCCCTTCAAAACCTGTAAGAACCTGTAAGTGTGAGTGCCTGATC 1680
DB |||||
DB 1621 TGTTCACCTTCCAGGCGCCCTTCAAAACCTGTAAGAACCTGTAAGTGTGAGTGCCTGATC 1680
QY 1681 AGAATTCGAGTGTGAGACCATAGACAGTTCGSCACTTACTTCTAATGCTTTTGGTGCATTTCA 1740
DB |||||
DB 1681 AGAATTCGAGTGTGAGACCATAGACAGTTCGSCACTTACTTCTAATGCTTTTGGTGCATTTCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATGCA 1800
DB |||||
DB 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATGCA 1800
QY 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGTGAGTGCCTTCACT 1860
DB |||||
DB 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGTGAGTGCCTTCACT 1860
QY 1861 TTTGGGAGCTTTGACGACATGCTGCTGGTGGGAGAAATGCGGTGCTGATGCTCAAT 1920
DB |||||
DB 1861 TTTGGGAGCTTTGACGACATGCTGCTGGTGGGAGAAATGCGGTGCTGATGCTCAAT 1920
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
DB |||||
DB 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
QY 1981 GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAAAGCAAAAGCTCCCAATTTCTAGC 2040
DB |||||
DB 1981 GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAAAGCAAAAGCTCCCAATTTCTAGC 2040
QY 2041 CTGAAAGTAAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB |||||
DB 2041 CTGAAAGTAAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB |||||
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTCTGCTTCTCTCATGATG 2220
DB |||||
DB 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTCTGCTTCTCTCATGATG 2220
QY 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGACAAAGGAGACCTGGAGAAATTTGG 2280
DB |||||
DB 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGACAAAGGAGACCTGGAGAAATTTGG 2280
QY 2281 GACTGCTCTATGTTAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
DB |||||
DB 2281 GACTGCTCTATGTTAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
QY 2341 CTGTCGCTTCTTGTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
DB |||||
DB 2341 CTGTCGCTTCTTGTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
QY 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
DB |||||
DB 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 ATCTTGTTCATCTCTCACTTTTAAGGAGGATCTGGTGGCTGAGAAAGCAAACTTACCTGCTC 2520
DB |||||

Db	2461	ATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGAAGCTTGAGAAAGCAAACTAGCTC	2522	
Qy	2521	TGGACAAGATCAAAACACCCCAAGCTTGAATGTCAATTAACCTGATGATGTCGAAAAACAG	2580	
Db	2521	TGGACAAGATCAAAACACCCCAAGCTTGAATGTCAATTAACCTGATGATGTCGAAAAACAG	2580	
Qy	2581	TCCTGTGACTCAACTCAAGCTTGGTTAACTTTACAGAGCTCCAGCATCACTTATGACCTG	2640	
Db	2581	TCCTGTGACTCAACTCAAGCTTGGTTAACTTTACAGAGCTCCAGCATCACTTATGACCTG	2640	
Qy	2641	CTCTCCAGTTCCTGGCCATCACCAGCTTATCCAGTGAATGAGAGCTGCCATCTTCTCTCT	2700	
Db	2641	CTCTCCAGTTCCTGGCCATCACCAGCTTATCCAGTGAATGAGAGCTGCCATCTTCTCTCT	2700	
Qy	2701	GTGGCAATTTGTCCCAATCTCTTAA	2724	
Db	2701	GTGGCAATTTGTCCCAATCTCTTAA	2724	
RESULT 10	AF061444	2724 bp	linear	PRI 17-NOV-1998
LOCUS	AF061444			
DEFINITION	Homo sapiens G protein-coupled receptor LGR5 (LGR5) mRNA, complete cds.			
ACCESSION	AF061444			
VERSION	AF061444.1	GI:3885471		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Hau, S.Y., Liang, S.G. and Haueh, A.J.			
TITLE	Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region			
JOURNAL	Mol. Endocrinol. 12 (12), 1830-1845 (1998)			
MEDLINE	93065210			
PUBMED	9849958			
REFERENCE	2 (bases 1 to 2724)			
AUTHORS	Hau, S.Y., Liang, S.G. and Haueh, A.J.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-APR-1998) Gyn/Ob, Stanford University, 300 Pasteur Dr., Stanford, CA 95305-5317, USA			
FEATURES	Location/Qualifiers			
source	1..2724			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
gene	1..2724			
	/gene="LGR5"			
CDS	1..2724			
	/gene="LGR5"			
	/notes="leucine-rich repeat containing G protein-coupled receptor; similar to glycoprotein hormone receptors; similar to FSH receptor, LH receptor and TSH receptor"			
	/codon_start=1			
	/product="G protein-coupled receptor LGR5"			
	/protein_id="AAC77911.1"			
	/db_xref="GI:3885472"			
	/translation="MPTSLRGLVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDG RMLLRVDCSLGSELPNSLQVLDLSNLLSMMNISLOSLPNPLFLHLELRLAGNAL TYTPKGAFTGSLKVLMLQNNLRHVPTAQLNRSLQSLDANHI SVVPPSPSG LHSRLHLLDNLATETIPQAFSLNLAQMTLANKIHI PDYAFGNLSVWVVLHL NNRHLSLKKCPGLHSLETLDDLNYNLDEFPPTAIRTLSNLKELGFHSNNIRSIPEKA FVNGPSLTIHFYDNPITQVGRSAFOHLPELRTLTNGASQITFEFDLTGTANLSLT LTQOQLSLRSLNLAQMTLQVLDLSNLLSMMNISLOSLPNPLFLHLELRLAGNAL TYTPKGAFTGSLKVLMLQNNLRHVPTAQLNRSLQSLDANHI SVVPPSPSG LHSRLHLLDNLATETIPQAFSLNLAQMTLANKIHI PDYAFGNLSVWVVLHL TGNHALQSLTSENPELKVIEPVAVQCAFVCENAYKISNOWNKGNSSMDLHLK XDGMFQAQDERLDEDFEDLKALHSVQCSPPFKCEKELDGLHRIIGWIT IAVLALCTNSPALTIVFRSPYISPIKLIGIVAAVNMGLTGVESSAVLHAGVDAPTFGSL ARHGAWWENGVCVHIVGSLISFASVSEVLLTALERGFVKYSKAFETKAPFSLML VITLLCALLALTAAPVLLGGSKYGASPLCLPLFPGEPTMGVWVALILNLSLCFLML			

QY 961 TTTCCTGATTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTTGGAGCACAGATC 1020
DB 961 TTTCCTGATTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTTGGAGCACAGATC 1020
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
DB 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
QY 1081 TACAACCTATTAGAAAGATTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
DB 1081 TACAACCTATTAGAAAGATTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
QY 1141 CTAAGACATATGAATCTACGAATTTAAAGTTTGAACATCTTTCCAGCAGTTCCTTAGCCTC 1200
DB 1141 CTAAGACATATGAATCTACGAATTTAAAGTTTGAACATCTTTCCAGCAGTTCCTTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAAGCTGACCTATCGTCCAACTCTGTCTGCTTTTCCCTATAACT 1320
DB 1261 TTGCCATCCCTAATAAAGCTGACCTATCGTCCAACTCTGTCTGCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAACAGGAAATCATGCTTACAGAGCTTTGATA 1380
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAACAGGAAATCATGCTTACAGAGCTTTGATA 1380
QY 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGTGT 1440
DB 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGTGT 1440
QY 1441 GCATTTGGAGTGTGAGAAATGCCCTTATAGAAATTTCTAATCAATGAAATAAAGGTGACAAAC 1500
DB 1441 GCATTTGGAGTGTGAGAAATGCCCTTATAGAAATTTCTAATCAATGAAATAAAGGTGACAAAC 1500
QY 1501 AGCAGTATGGAGACCTTTCATAGAAAGTGTGGAATGTTTCAGGCTCAAGATGAACGT 1560
DB 1501 AGCAGTATGGAGACCTTTCATAGAAAGTGTGGAATGTTTCAGGCTCAAGATGAACGT 1560
QY 1561 GACCTTTGAAGATTTCCTGCTTGACCTTTCAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1561 GACCTTTGAAGATTTCCTGCTTGACCTTTCAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGGCTGGCTGATC 1680
QY 1681 AGAATTTGGAGTGTGGAACCATAGCAGTCTGGCAGCTTACTTGAATGCTTTGGTGAATCTCA 1740
DB 1681 AGAATTTGGAGTGTGGAACCATAGCAGTCTGGCAGCTTACTTGAATGCTTTGGTGAATCTCA 1740
QY 1741 ACAGATTTTCAGATCCCTCTGTACATTTCCCCCAATTAATTTAACTGTTAACTGGGTTCATCGCA 1800
DB 1741 ACAGATTTTCAGATCCCTCTGTACATTTCCCCCAATTAATTTAACTGTTAACTGGGTTCATCGCA 1800
QY 1801 GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGTGCTGGCTGGTGTGATGCGTTCAC 1860
DB 1801 GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGTGCTGGCTGGTGTGATGCGTTCAC 1860
QY 1861 TTTCGACGCTTGCACGACATGGTCCCTGCTGGGAGAAATGGGTTGGTGGCCATGCTCATT 1920
DB 1861 TTTCGACGCTTGCACGACATGGTCCCTGCTGGGAGAAATGGGTTGGTGGCCATGCTCATT 1920
QY 1921 GGTGTTTTGTCATTTTGTCTTCAGAACTCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
DB 1921 GGTGTTTTGTCATTTTGTCTTCAGAACTCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
QY 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAAATCAITTTTGTCTGTGCCCTGTGCTGGCCTTGACCATGGCCGAGTTCCTCCCTG 2100

DB 2041 CTGAAAGTAAATCAITTTTGTCTGTGCCCTGTGCTGCTTACCATGGCCGAGTTCCTCCCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGGCTTTTGGGGAGCC 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGGCTTTTGGGGAGCC 2160
QY 2161 AGCACCATGGCTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGGCTTCTCATGATG 2220
DB 2161 AGCACCATGGCTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGGCTTCTCATGATG 2220
QY 2221 ACCATTTGCTACCAAGCTCTACTGCAATTTTGACAAGGAGAGACCTGGAGAAATATTTGG 2280
DB 2221 ACCATTTGCTACCAAGCTCTACTGCAATTTTGACAAGGAGAGACCTGGAGAAATATTTGG 2280
QY 2281 GACTGTCTATTTGTPAAACACATTTGCCCTGTGTCTTTCACCAACTGATCCTTAAACTGTC 2340
DB 2281 GACTGTCTATTTGTPAAACACATTTGCCCTGTGTCTTTCACCAACTGATCCTTAAACTGTC 2340
QY 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCCTTACATTTATCAGTCTCGAAGTA 2400
DB 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCCTTACATTTATCAGTCTCGAAGTA 2400
QY 2401 ATTAAGTTTATCTTCTGCTGCTAGTCCCACTCTCTGCTGCTCTCAATCCCCTTCTCTAC 2460
DB 2401 ATTAAGTTTATCTTCTGCTGCTAGTCCCACTCTCTGCTGCTCTCAATCCCCTTCTCTAC 2460
QY 2461 ATCTTTGTTCAATCTCACTTTTAAGGAGATCTGTGAGCCTGTGAGAAAGCAACCTACGTC 2520
DB 2461 ATCTTTGTTCAATCTCACTTTTAAGGAGATCTGTGAGCCTGTGAGAAAGCAACCTACGTC 2520
QY 2521 TGGACAGATCAAAACACCCCAAGCTTGTGATGTCATTAACCTGATGATGTCGAAAAACAG 2580
DB 2521 TGGACAGATCAAAACACCCCAAGCTTGTGATGTCATTAACCTGATGATGTCGAAAAACAG 2580
QY 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCACCTTATGACCTG 2640
DB 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCACCTTATGACCTG 2640
QY 2641 CCTCCCAGTTCCTGTCACCACTGATTCACAGTGTATCCAGTGTGAGAGTGCATCTTCTCTCT 2700
DB 2641 CCTCCCAGTTCCTGTCACCACTGATTCACAGTGTATCCAGTGTGAGAGTGCATCTTCTCTCT 2700
QY 2701 GTGGCAATTTGCCCATGTCCTAA 2724
DB 2701 GTGGCAATTTGCCCATGTCCTAA 2724

RESULT 11
BD135244

LOCUS BD135244 2082 bp DNA linear PAT 18-SEP-2002
DEFINITION Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain.
ACCESSION BD135244
VERSION BD135244.1 GI:23230189
KEYWORDS JP 2002507406-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2082)
AUTHORS Hsueh,A.J.W.; Hsu,S.Y.; Liang,S.G. and Spe,P.J.V.D.
TITLE Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain
JOURNAL Patent: JP 2002507406-A 2 12-MAR-2002;
THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,
AKZO NOBEL NV
COMMENT OS Homo sapiens (human)
PN JP 2002507406-A/2
PD 12-MAR-2002
PF 25-MAR-1999 JP 2000537903
PR 26-MAR-1998 US 60/079501
PI AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 14:17:43 ; Search time 1408 Seconds
(without alignments)
11452.681 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 2724

Sequence: 1 atggacacctccggctcgg.....cattgtcccatgtctctaa 2724

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2724	100.0	2724	3 Aa30770	Aa30770 Human G p
2	2724	100.0	2724	8 Abz42816	Abz42816 Human G p
3	2724	100.0	2724	10 Adc22782	Adc22782 Human G p
4	2724	100.0	2724	10 Adh14255	Adh14255 Human HG3
5	2724	100.0	2724	12 Adl32985	Adl32985 Human G p
6	2724	100.0	2724	12 Ado29923	Ado29923 Human GPC
7	2724	100.0	2880	11 Adn39796	Adn39796 Cancer/an
8	2724	100.0	2880	13 Adq80249	Adq80249 G protein
9	2724	100.0	2973	13 Adr67869	Adr67869 Human HG3
10	2724	100.0	3032	11 Adf39627	Adf39627 Cancer/an
11	2724	100.0	3032	11 Adn39530	Adn39530 Cancer/an
12	2724	100.0	4558	2 Aax23980	Aax23980 Human HG3
13	2724	100.0	4570	12 Adl12472	Adl12472 Human etc
14	2670	98.0	3438	10 Adf70582	Adf70582 Orphan re
15	2622	96.3	2724	3 Aa30779	Aa30779 DNA encod
16	2622	96.3	2724	10 Adc22796	Adc22796 Human G p
17	2622	96.3	2724	10 Adh14269	Adh14269 Mutated h
18	2603	95.6	3297	2 Aax23981	Aax23981 Human HG3
19	1864	68.4	2651	10 Adb80463	Adb80463 Ovarian c
20	1864	68.4	2651	11 Adn39165	Adn39165 Cancer/an

21	1864	68.4	2651	11	ADN39795
22	1842	67.6	2082	2	Aa25344
23	699	25.7	723	8	ACA04772
24	641	23.5	1790	6	ABQ5077
25	606	22.2	606	2	Aa240459
26	72	2.6	2636	5	ABA09697
27	60	2.2	60	6	ABN47039
28	41	1.5	2724	12	ADO30213
29	41	1.5	3098	13	ADR67870
30	30	1.1	30	12	ADI32988
31	25	0.9	25	3	AAA30768
32	25	0.9	25	10	ADC22780
33	25	0.9	25	10	ADH14253
34	24	0.9	30	3	AAA30767
35	24	0.9	30	10	ADC22779
36	24	0.9	30	10	ADH14252
37	24	0.9	2901	6	AAI67921
38	24	0.9	2901	11	ADN02241
39	24	0.9	2901	12	ADK19406
40	24	0.9	3637	6	AAI67920
41	24	0.9	3637	11	ADN02239
42	24	0.9	3637	12	ADK19404
43	22	0.8	22	2	AAX23973
44	22	0.8	22	2	AAX23974
45	22	0.8	22	2	AAX23972

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
AC AAA30770;
XX
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.

KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.
XX Homo sapiens.
XX WO200022129-A1.
XX PD 20-APR-2000.
XX PF 12-OCT-1999; 99WO-US023938.
XX PR 13-OCT-1998; 98US-00170496.
XX (AREN-) ARENA PHARM INC.
XX Behan DP, Chalmers DT, Liaw CW;
XX WPI: 2000-329165/28.
XX P-PSDB; AAY90682.

Non-endogenous constitutively activated human G protein-coupled receptors, useful for identifying agonists for use as pharmaceutical agents.

Example 1; Page 315-317; 341pp; English.

The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-CC AAY90677 and AAY90683-Y90687) and to DNA encoding them (AAA30709-A30743 and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, x,

Qy	61	GGGGCAGCTCTCCAGGCTCTGGTGTGGTGGGGCTGCCACACATGTCTATGC	120
Db			
Qy	61	GGGGCAGCTCTCCAGGCTCTGGTGTGGTGGGGCTGCCACACATGTCTATGC	120
Db			
Qy	121	GAGCCGACGGCAGGATGTGCTCAGGSGTGACTCTCCGACCTGGGGCTCTCGAGCTG	180
Db			
Qy	121	GAGCCGACGGCAGGATGTGCTCAGGSGTGACTCTCCGACCTGGGGCTCTCGAGCTG	180
Db			
Qy	181	CTTCTCAACCTCAGGGTCTTCACTCTCTCCTPACCTAGACCTCAGTATGAACAAATCAGTCA	240
Db			
Qy	181	CTTCTCAACCTCAGGGTCTTCACTCTCTCCTPACCTAGACCTCAGTATGAACAAATCAGTCA	240
Db			
Qy	241	CTGCTCCGAATCCCTCCGACCTCCGCTTCCCTGGAGGATTCAGCTCTGGCGGAAC	300
Db			
Qy	241	CTGCTCCGAATCCCTCCGACCTCCGCTTCCCTGGAGGATTCAGCTCTGGCGGAAC	300
Db			
Qy	301	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTCTTTATG	360
Db			
Qy	301	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTCTTTATG	360
Db			
Qy	361	CTGCAGAAATACAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT	420
Db			
Qy	361	CTGCAGAAATACAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT	420
Db			
Qy	421	CAATCCCTGCGTCTGGATGCTAAACACATCAGTATGTGCCCAAGCTGTTTCAAGTGGC	480
Db			
Qy	421	CAATCCCTGCGTCTGGATGCTAAACACATCAGTATGTGCCCAAGCTGTTTCAAGTGGC	480
Db			
Qy	481	CTGCATTCCTCAGGACACTGTGGCTGGATGACAAATGGCTTAAACAGAAATCCCGTCCAG	540
Db			
Qy	481	CTGCATTCCTCAGGACACTGTGGCTGGATGACAAATGGCTTAAACAGAAATCCCGTCCAG	540
Db			
Qy	541	GCTTTTGAAGTTTATCGGCAATGCAAGCCTGCTGGCCCTGAAACAAATATACACAC	600
Db			
Qy	541	GCTTTTGAAGTTTATCGGCAATGCAAGCCTGCTGGCCCTGAAACAAATATACACAC	600
Db			
Qy	601	ATACAGACTATGCTTTGGAACTCTCCAGCTGTGTAGTCTACATCTCCATACCAAT	660
Db			
Qy	601	ATACAGACTATGCTTTGGAACTCTCCAGCTGTGTAGTCTACATCTCCATACCAAT	660
Db			
Qy	661	AGAACTCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Db			
Qy	661	AGAACTCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Db			
Qy	721	TTAAATTAACAATAAATCTTGATGAATTCGCCACTGCAATTTAGGACACTCTCCAAACCTTAA	780
Db			
Qy	721	TTAAATTAACAATAAATCTTGATGAATTCGCCACTGCAATTTAGGACACTCTCCAAACCTTAA	780
Db			
Qy	781	GAACTAGGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db			
Qy	781	GAACTAGGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db			
Qy	841	CCCTCTCTTTATTAACAATATTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT	900
Db			
Qy	841	CCCTCTCTTTATTAACAATATTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT	900
Db			
Qy	901	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTTCACAATTAACCTGAA	960
Db			
Qy	901	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTTCACAATTAACCTGAA	960
Db			
Qy	961	TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db			
Qy	961	TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db			
Qy	1021	TCACTCTCTTCCAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGTCT	1080
Db			
Qy	1021	TCACTCTCTTCCAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGTCT	1080
Db			
Qy	1081	TACAACCTATTTAGAAAGTTTACCCAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC	1140
Db			
Qy	1081	TACAACCTATTTAGAAAGTTTACCCAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC	1140
Db			

Qy	1141	CTAAGACATTAATGAAATCTACGAAATTTAAAGTTTGACACTTTTCAGACAGTTGCTTTAGCCTC	1200
Db			
Qy	1141	CTAAGACATTAATGAAATCTACGAAATTTAAAGTTTGACACTTTTCAGACAGTTGCTTTAGCCTC	1200
Db			
Qy	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT	1260
Db			
Qy	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT	1260
Db			
Qy	1261	TTGCGCATCCCTTAATAAGAGTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATTAACCT	1320
Db			
Qy	1261	TTGCGCATCCCTTAATAAGAGTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATTAACCT	1320
Db			
Qy	1321	GGGTTCATAGTGTAACTCACTTAAATTAACAGAAATCATGCTTTACAGAGCTTGATA	1380
Db			
Qy	1321	GGGTTCATAGTGTAACTCACTTAAATTAACAGAAATCATGCTTTACAGAGCTTGATA	1380
Db			
Qy	1381	TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGTTTACAGTGTCTGT	1440
Db			
Qy	1381	TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGTTTACAGTGTCTGT	1440
Db			
Qy	1441	GCATTTGGAGTGTGTGAGAAATGCTTAAATTAATCAATGGAATTAAGGTGACAAC	1500
Db			
Qy	1441	GCATTTGGAGTGTGTGAGAAATGCTTAAATTAATCAATGGAATTAAGGTGACAAC	1500
Db			
Qy	1501	AGCAGTATGACGACCTTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAAAGT	1560
Db			
Qy	1501	AGCAGTATGACGACCTTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAAAGT	1560
Db			
Qy	1561	GACCTTGAAGATTTCTGCTTGAATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAG	1620
Db			
Qy	1561	GACCTTGAAGATTTCTGCTTGAATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAG	1620
Db			
Qy	1621	TGTTCACTTTCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGTGATGGCTGCTGATC	1680
Db			
Qy	1621	TGTTCACTTTCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGTGATGGCTGCTGATC	1680
Db			
Qy	1681	AGAAATGGAGTGTGGAACATAGCAGATTTCTGGAATTTCTGTAATGCTTTGGTGACTTCA	1740
Db			
Qy	1681	AGAAATGGAGTGTGGAACATAGCAGATTTCTGGAATTTCTGTAATGCTTTGGTGACTTCA	1740
Db			
Qy	1741	ACAGTTTTCAGATCCCTCTGTGTAATTTCCGCCAATTAACCTGTTAATTTGGGGTCAATGCA	1800
Db			
Qy	1741	ACAGTTTTCAGATCCCTCTGTGTAATTTCCGCCAATTAACCTGTTAATTTGGGGTCAATGCA	1800
Db			
Qy	1801	GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCTGTGGAACACCTGTGATGGCTGCTGATC	1860
Db			
Qy	1801	GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCTGTGGAACACCTGTGATGGCTGCTGATC	1860
Db			
Qy	1861	TTTGGCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGCATGTCAAT	1920
Db			
Qy	1861	TTTGGCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGCATGTCAAT	1920
Db			
Qy	1921	GGTTTTTGTCCATTTTGTTCAGAAATCATCTGTTTCTGCTTACTCTGGAGCCCTG	1980
Db			
Qy	1921	GGTTTTTGTCCATTTTGTTCAGAAATCATCTGTTTCTGCTTACTCTGGAGCCCTG	1980
Db			
Qy	1981	GAGCGTGGTCTCTGTGGAATATTTGCAAAATTTGAAACGAAAGCTTCAATTTTCTAGC	2040
Db			
Qy	1981	GAGCGTGGTCTCTGTGGAATATTTGCAAAATTTGAAACGAAAGCTTCAATTTTCTAGC	2040
Db			
Qy	2041	CTGAAAGTAAATCAATTTTGTCTGTGCTGCTGGCTTGACCATGGCGCAGTTTCCCTG	2100
Db			
Qy	2041	CTGAAAGTAAATCAATTTTGTCTGTGCTGCTGGCTTGACCATGGCGCAGTTTCCCTG	2100
Db			
Qy	2101	CTGGGTGGCAGCAAGTATGGGCTCTCCCTCTCTGCTGCTGCTTTTGGGGAGGCC	2160
Db			
Qy	2101	CTGGGTGGCAGCAAGTATGGGCTCTCCCTCTCTGCTGCTGCTTTTGGGGAGGCC	2160
Db			
Qy	2161	AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATCCCTTGTCTCTCATGATG	2220
Db			
Qy	2161	AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATCCCTTGTCTCTCATGATG	2220
Db			
Qy	2221	ACCATTGCTACACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTTGG	2280
Db			

Qy		661	AGAAATCACTCCCTGGGAAAAGAAATGCTTTGAATGGGTCTCCACAGCCTTAGAGACTTTAGAT	720
Db		661	AGAATCCACTCCCTGGGAAAAGAAATGCTTTGAATGGGTCTCCACAGCCTTAGAGACTTTAGAT	720
Qy		721	TTTAAATTACAAATAACCTTTGATGAATTCGCCACCTGCATTAAGGACACTCTCCAACCTTAAA	780
Db		721	TTTAAATTACAAATAACCTTTGATGAATTCGCCACCTGCATTAAGGACACTCTCCAACCTTAAA	780
Qy		781	GAACTAGGATTTCATATAGCAACAATATCAGGTGCGATCCTTGAGAAAGCAITTTGTAGGCAAC	840
Db		781	GAACTAGGATTTCATATAGCAACAATATCAGGTGCGATCCTTGAGAAAGCAITTTGTAGGCAAC	840
Qy		841	CCTTCTCTTAATACAAATATCTATATGACAAATCCCACATCCCAATCCCAATCCCAATCCCAAT	900
Db		841	CCTTCTCTTAATACAAATATCTATATGACAAATCCCACATCCCAATCCCAATCCCAATCCCAAT	900
Qy		901	TTTTCAAACATTTACTCGAAGCTTAAGAACACTGACCTGGAATGGTGGCTCACAAATAACTGAA	960
Db		901	TTTTCAAACATTTACTCGAAGCTTAAGAACACTGACCTGGAATGGTGGCTCACAAATAACTGAA	960
Qy		961	TTTCTCTGATTTAACTCGAACTGCAAACTCGAGAGTCTGACTTTTAACTGAGGACACAGATC	1020
Db		961	TTTCTCTGATTTAACTCGAACTGCAAACTCGAGAGTCTGACTTTTAACTGAGGACACAGATC	1020
Qy		1021	TCACT	1080
Db		1021	TCACT	1080
Qy		1081	TAGAACCTATTTAGAGATTTTACCAGTTTTTTCAGTCTGCCAAGCTTTCAGAAATTTGAC	1140
Db		1081	TAGAACCTATTTAGAGATTTTACCAGTTTTTTCAGTCTGCCAAGCTTTCAGAAATTTGAC	1140
Qy		1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTTCTTAGGCTC	1200
Db		1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTTCTTAGGCTC	1200
Qy		1201	CGATCGCTGAAATTTGGCTTGGAACAAAAATGCTATTATCCOCCAAATGCAATTTTCCACT	1260
Db		1201	CGATCGCTGAAATTTGGCTTGGAACAAAAATGCTATTATCCOCCAAATGCAATTTTCCACT	1260
Qy		1261	TTGGCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTCTGCTGGCTTTTCTCTATTAACCT	1320
Db		1261	TTGGCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTCTGCTGGCTTTTCTCTATTAACCT	1320
Qy		1321	GGGTTACATGGTTTAACTCACTTAAATTTAACAGGAAATCATGCTCTTACAGAGCTTGATA	1380
Db		1321	GGGTTACATGGTTTAACTCACTTAAATTTAACAGGAAATCATGCTCTTACAGAGCTTGATA	1380
Qy		1381	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTCTGT	1440
Db		1381	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTCTGT	1440
Qy		1441	GCATTTGGAGTGTGTGAAATGCCCTATAGATTTCTAATCAATGGAAATAAGGTGACAAC	1500
Db		1441	GCATTTGGAGTGTGTGAAATGCCCTATAGATTTCTAATCAATGGAAATAAGGTGACAAC	1500
Qy		1501	AGCAGTATGACAGCACTTCTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db		1501	AGCAGTATGACAGCACTTCTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Qy		1561	GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGAC	1620
Db		1561	GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGAC	1620
Qy		1621	TGTTTCACTTCCCGAGGCCCTTCAAAACCTGTGAAACCTGTGATGGCTGGCTGATC	1680
Db		1621	TGTTTCACTTCCCGAGGCCCTTCAAAACCTGTGAAACCTGTGATGGCTGGCTGATC	1680
Qy		1681	AGAAATCGAGTGTGGAACATAGCAGTTCTGSCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db		1681	AGAAATCGAGTGTGGAACATAGCAGTTCTGSCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Qy		1741	ACAGTTTTTCAGATCCCCCTCTGTACATTTTCCCCCAATTAACCTGTTAAATGGGGTTCATCGCA	1800

Db	1741		ACAGT	TTT	CAGAT	CCCT	CTGT	ACAT	TT	CCCC	CAAT	TAA	ACT	GT	TAAT	TGGGG	TCAT	CGCA	1800							
Qy	1801		GCA	GT	GAA	CA	TG	CT	CAC	GGA	GT	CT	CAG	TG	CCG	TG	CT	GGT	TGGA	TGG	TTC	CACT	1860			
Db	1801		GCA	GT	GAA	CA	TG	CT	CAC	GGA	GT	CT	CAG	TG	CCG	TG	CT	GGT	TGGA	TGG	TTC	CACT	1860			
Qy	1861		TTT	GG	CAG	CT	TT	GCA	CGA	CA	TG	TG	CT	GGT	GGG	AGAA	TGG	GGT	TGG	TCC	CA	TG	CAAT	1920		
Db	1861		TTT	GG	CAG	CT	TT	GCA	CGA	CA	TG	TG	CT	GGT	GGG	AGAA	TGG	GGT	TGG	TCC	CA	TG	CAAT	1920		
Qy	1921		GG	TTTT	TT	GT	CC	TT	TG	CT	T	CAG	AA	CA	T	CT	GT	TT	T	C	T	GT	CA	CT	1980	
Db	1921		GG	TTTT	TT	GT	CC	TT	TG	CT	T	CAG	AA	CA	T	CT	GT	TT	T	C	T	GT	CA	CT	1980	
Qy	1981		GAG	CG	TGG	GT	CT	CT	G	TG	AA	AA	T	T	T	G	AA	A	TT	T	G	AA	A	TT	2040	
Db	1981		GAG	CG	TGG	GT	CT	CT	G	TG	AA	AA	T	T	T	G	AA	A	TT	T	G	AA	A	TT	2040	
Qy	2041		CT	GAA	GT	AA	T	CA	TT	T	T	G	CT	CT	G	CG	CT	T	G	A	CA	T	G	CG	CT	2100
Db	2041		CT	GAA	GT	AA	T	CA	TT	T	T	G	CT	CT	G	CG	CT	T	G	A	CA	T	G	CG	CT	2100
Qy	2101		CT	GG	T	GG	C	A	G	A	G	T	A	T	G	C	G	C	T	C	T	C	T	C	T	2160
Db	2101		CT	GG	T	GG	C	A	G	A	G	T	A	T	G	C	G	C	T	C	T	C	T	C	T	2160
Qy	2161		AG	CA	CA	T	G	GG	T	CG	CT	C	A	T	G	CT	C	A	T	G	CT	C	A	T	G	2220
Db	2161		AG	CA	CA	T	G	GG	T	CG	CT	C	A	T	G	CT	C	A	T	G	CT	C	A	T	G	2220
Qy	2221		ACC	AT	T	GC	CT	T	A	C	CA	A	G	C	T	T	T	G	G	A	A	G	G	A	A	2280
Db	2221		ACC	AT	T	GC	CT	T	A	C	CA	A	G	C	T	T	T	G	G	A	A	G	G	A	A	2280
Qy	2281		GA	CT	CT	AT	AG	T	AA	CA	CA	CA	T	T	GC	CT	CT	CA	CA	AT	GC	AT	CT	CA	AT	2340
Db	2281		GA	CT	CT	AT	AG	T	AA	CA	CA	CA	T	T	GC	CT	CT	CA	CA	AT	GC	AT	CT	CA	AT	2340
Qy	2341		CC	T	G	T	G	CT	TT	C	T	T	T	A	A	A	C	CT	T	A	C	AT	T	T	A	2400
Db	2341		CC	T	G	T	G	CT	TT	C	T	T	T	A	A	A	C	CT	T	A	C	AT	T	T	A	2400
Qy	2401		AT	T	A	A	G	T	T	A	T	C	CT	T	CG	AT	CC	CA	T	CT	CA	AT	CC	CT	CT	2460
Db	2401		AT	T	A	A	G	T	T	A	T	C	CT	T	CG	AT	CC	CA	T	CT	CA	AT	CC	CT	CT	2460
Qy	2461		AT	CT	T	GT	T	CA	AT	CT	TT	T	AG	G	AG	T	CT	GT	G	AG	C	CT	G	A	A	2520
Db	2461		AT	CT	T	GT	T	CA	AT	CT	TT	T	AG	G	AG	T	CT	GT	G	AG	C	CT	G	A	A	2520
Qy	2521		TG	GA	CA	AG	T	CA	AA	CA	CC	CA	G</													

RESULT 4
ADH14255
ID ADH14255 standard; cDNA; 2724 BP.
XX
AC ADH14255;

Qy	1681	AGAAATGGAGTGTGCACCATAGCAGTTCTGSCACTTA	CTTGTAATGCTTTGGTGCACTTCA	1740
Db	1681	AGAAATGGAGTGTGCACCATAGCAGTTCTGSCACTTA	CTTGTAATGCTTTGGTGCACTTCA	1740
Qy	1741	ACAGTTTTAGATCCCTCTGTACATTTCGCCCATTA	AAGCTGTTAAATGGGGTCATCGCA	1800

Db 61 GGGGGCAGCTCTCCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCTATTGC 120
Qy 121 GAGCCGACGCGAGGATGTTGCTCAGGGTGAAGTGTCTCGACCTGGGGCTCTCGGAGCTG 180
Db 121 GAGCCGACGCGAGGATGTTGCTCAGGGTGAAGTGTCTCGACCTGGGGCTCTCGGAGCTG 180
Qy 181 CTTTCCAACTCAGCGCTTCCACTCCTACCTAGACCTCAGTATGAACACATCAGTCAG 240
Db 181 CTTTCCAACTCAGCGCTTCCACTCCTACCTAGACCTCAGTATGAACACATCAGTCAG 240
Qy 241 CTGCTCCGAATCCCTCGCCAGTCTCCGCTTCTCGAGGAGTACGTCTTGCGGGAAC 300
Db 241 CTGCTCCGAATCCCTCGCCAGTCTCCGCTTCTCGAGGAGTACGTCTTGCGGGAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGAGCAATTCAGTGGCTTTACAGTCTTAAAGTTCATTG 360
Db 301 GCTCTGACATACATTTCCCAAGGAGCAATTCAGTGGCTTTACAGTCTTAAAGTTCATTG 360
Qy 361 CTGCAGAAATAACAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAGCCTT 420
Db 361 CTGCAGAAATAACAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAGCCTT 420
Qy 421 CAATCCCTGGTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
Db 421 CAATCCCTGGTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
Qy 481 CTGCATTTCCCTGAGGCACCTGTGGCTGGATGACAAATGGCTTAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTTCCCTGAGGCACCTGTGGCTGGATGACAAATGGCTTAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCAATTCGAGCAATGACCAATGACCTTGGCCCTGAAACACCA 600
Db 541 GCTTTTGAAGTTTATCGGCAATTCGAGCAATGACCAATGACCTTGGCCCTGAAACACCA 600
Qy 601 ATACACAGATATGCCCTTTGGAACCTCTCCAGCTTGGTAGTCTTACATCTCCATACCAAT 660
Db 601 ATACACAGATATGCCCTTTGGAACCTCTCCAGCTTGGTAGTCTTACATCTCCATACCAAT 660
Qy 661 AGAATCCATCCCTGCGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCATCCCTGCGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAAATTACAAATAACCTTGAATTTCCCACTGCAATTAGGACACTCTCCAACCTTAA 780
Db 721 TTAAATTACAAATAACCTTGAATTTCCCACTGCAATTAGGACACTCTCCAACCTTAA 780
Qy 781 GAACTAGGATTTATAGCAACAAATATCAGTGGATACCTGAGAAAGCAATTTGAGGCAAC 840
Db 781 GAACTAGGATTTATAGCAACAAATATCAGTGGATACCTGAGAAAGCAATTTGAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATTTCTATGACAAATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATTTCTATGACAAATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 960
Db 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 960
Qy 961 TTTTCCGTGATTTAAGTGAACCTGAAACCTGGAGGCTGACCTTAACTGGAGCAGATC 1020
Db 961 TTTTCCGTGATTTAAGTGAACCTGAAACCTGGAGGCTGACCTTAACTGGAGCAGATC 1020
Qy 1021 TCATCTCTTCCCAACCGCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGTCT 1080
Db 1021 TCATCTCTTCCCAACCGCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAAGAAATTTACCGAGTTTTCAGTCTGCAAAAGCTTCAGAAAATGAC 1140
Db 1081 TACAACCTATTAAGAAATTTACCGAGTTTTCAGTCTGCAAAAGCTTCAGAAAATGAC 1140
Qy 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTGCTAGCCTC 1200
Db 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTGCTAGCCTC 1200

Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTTATTCACCCCAATGCAATTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTTATTCACCCCAATGCAATTTCCACT 1260
Qy 1261 TTTGCCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTCTATAACT 1320
Db 1261 TTTGCCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTCTATAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGAAATCATGCCCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGAAATCATGCCCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTCTGT 1440
Db 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTCTGT 1440
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCCCTTATAGATTTCTTAATCAATGGAATTAAGGTGCAAC 1500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCCCTTATAGATTTCTTAATCAATGGAATTAAGGTGCAAC 1500
Qy 1501 AGCAGTATGGACACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGGACACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
Db 1621 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680
Qy 1681 AGAATTTGAGTGTGACCATAGCAGTCTTGGCACTTACTTTGTAATGCTTTTGGTGACTTCA 1740
Db 1681 AGAATTTGAGTGTGACCATAGCAGTCTTGGCACTTACTTTGTAATGCTTTTGGTGACTTCA 1740
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCGCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCGCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Qy 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCGTGTGGCTGGTGTGATGGCTGCTTCACT 1860
Db 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCGTGTGGCTGGTGTGATGGCTGCTTCACT 1860
Qy 1861 TTTGGCAGCTTTGACAGCATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
Db 1861 TTTGGCAGCTTTGACAGCATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTCTGAAAAATTTGAAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTCTGAAAAATTTGAAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAACTCATTTTGTCTGTGCCCTGTGGCTTGAACCATGCGCAGTTCCTCTG 2100
Db 2041 CTGAAAGTAACTCATTTTGTCTGTGCCCTGTGGCTTGAACCATGCGCAGTTCCTCTG 2100
Qy 2101 CTGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTTGTGCTTTTGGGGAGGCC 2160
Db 2101 CTGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTTGTGCTTTTGGGGAGGCC 2160
Qy 2161 AGCAGCATGGGTACATGGTCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 2161 AGCAGCATGGGTACATGGTCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Qy 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTGGAAGGAGGAGACCTGGAGAAATTTGG 2280
Db 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTGGAAGGAGGAGACCTGGAGAAATTTGG 2280

QY 361 CTGCGAATAATCAGCTAAGACACGTAACCCACAGAAAGCTCTGCGAGAAATTTGGGAGCCCTT 420
DB 361 CTGCGAATAATCAGCTAAGACACGTAACCCACAGAAAGCTCTGCGAGAAATTTGGGAGCCCTT 420
QY 421 CAATCCCTCGCTGCTGATGCTAAGACACATCAGCTATGTCGCCCCCAAGCTGTTTCAGTGGC 480
DB 421 CAATCCCTCGCTGCTGATGCTAAGACACATCAGCTATGTCGCCCCCAAGCTGTTTCAGTGGC 480
QY 481 CTGCAATCCCTGAGGACCTGTGGCTGATGATCAATGCGTTAAGCAGAAATCCCGTCCAG 540
DB 481 CTGCAATCCCTGAGGACCTGTGGCTGATGATCAATGCGTTAAGCAGAAATCCCGTCCAG 540
QY 541 GCTTTTAGAAGTTTATCGGCATTCGAGCATGCAAGCATGACCTTGGCCCTGAAACAAATACACAC 600
DB 541 GCTTTTAGAAGTTTATCGGCATTCGAGCATGCAAGCATGACCTTGGCCCTGAAACAAATACACAC 600
QY 601 ATACACAGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660
DB 601 ATACACAGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
QY 721 TTAATAATCAATAACCTTGATGAATTTCCCGACTGCAATTTAGGACACTCTCCAACTTAAA 780
DB 721 TTAATAATCAATAACCTTGATGAATTTCCCGACTGCAATTTAGGACACTCTCCAACTTAAA 780
QY 781 GAATAGGATTTTATAGCAACATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 781 GAATAGGATTTTATAGCAACATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 841 CTTCTCTTTTATACATAATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT 900
DB 841 CTTCTCTTTTATACATAATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT 900
QY 901 TTTCAACATTTACTGAACTAAGAACACTGACTCTGAAATGGTGGCTTCAAAATAACTGAA 960
DB 901 TTTCAACATTTACTGAACTAAGAACACTGACTCTGAAATGGTGGCTTCAAAATAACTGAA 960
QY 961 TTTCTGATTTTAACTGGAATTCGAACTGGAAGCTGAGAGCTGACATTTAACTGGAGCAGATC 1020
DB 961 TTTCTGATTTTAACTGGAATTCGAACTGGAAGCTGAGAGCTGACATTTAACTGGAGCAGATC 1020
QY 1021 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGCT 1080
DB 1021 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGCT 1080
QY 1081 TACAACCTTATAGAGATTTACCAGTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
DB 1081 TACAACCTTATAGAGATTTACCAGTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
QY 1141 CTAAGCATAATGAAATCTACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1200
DB 1141 CTAAGCATAATGAAATCTACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1200
QY 1201 CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTATTCACCCCAATGCAATTTCCACT 1260
DB 1201 CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTATTCACCCCAATGCAATTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAGCTGAGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
DB 1261 TTGCCATCCCTAATAAGCTGAGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTTATGCTTTACCAGTGTCT 1440
DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTTATGCTTTACCAGTGTCT 1440

QY 1441 GCATTTGGAGTGTGAGAAATGCCCTATAAGATTTTCTAATCAATGGAATTAAGGTGACAC 1500
DB 1441 GCATTTGGAGTGTGAGAAATGCCCTATAAGATTTTCTAATCAATGGAATTAAGGTGACAC 1500
QY 1501 AGCAGTATGAGACGACCTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACCT 1560
DB 1501 AGCAGTATGAGACGACCTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACCT 1560
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY 1621 TGTTTCACTTCCCAAGGCCCTTCAACCCCTGTGAAACCTCTGCTTGAATGGCTGGCTGATC 1680
DB 1621 TGTTTCACTTCCCAAGGCCCTTCAACCCCTGTGAAACCTCTGCTTGAATGGCTGGCTGATC 1680
QY 1681 AGAATTTGAGTGTGAGACCATAGCAGTTCTGGCACTTCTGTAATGCTTTTGGTGAATTTCA 1740
DB 1681 AGAATTTGAGTGTGAGACCATAGCAGTTCTGGCACTTCTGTAATGCTTTTGGTGAATTTCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGTATCAATTTCCGCCCATTTAAATGGGGTCAATCGCA 1800
DB 1741 ACAGTTTTCAGATCCCTCTGTATCAATTTCCGCCCATTTAAATGGGGTCAATCGCA 1800
QY 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGCTGGTGTGGATGGTTCAT 1860
DB 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGCTGGTGTGGATGGTTCAT 1860
QY 1861 TTTGGCAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATT 1920
DB 1861 TTTGGCAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATT 1920
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTTGGCAGCCCTG 1980
DB 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTTGGCAGCCCTG 1980
QY 1981 GAGCTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAAACGAAAGTCTCATTTCTAGC 2040
DB 1981 GAGCTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAAACGAAAGTCTCATTTCTAGC 2040
QY 2041 CTGAAAGTAATCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGAAAGTAATCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTTGGCGAGGCC 2160
DB 2101 CTGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTTGGCGAGGCC 2160
QY 2161 AGCACATGGGCTACATGGTGTCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
DB 2161 AGCACATGGGCTACATGGTGTCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATTTGG 2280
DB 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATTTGG 2280
QY 2281 GACTGCTCTATGGTAAACACATTTGCTGCTTCTTCCAACTGCAATCTCTAAACTGC 2340
DB 2281 GACTGCTCTATGGTAAACACATTTGCTGCTTCTTCCAACTGCAATCTCTAAACTGC 2340
QY 2341 CCTGTGGCTTCTTGTCTCTTCTCTCTTAAACCTTTACATTTATCAGTCTGAGTA 2400
DB 2341 CCTGTGGCTTCTTGTCTCTTCTCTCTTAAACCTTTACATTTATCAGTCTGAGTA 2400
QY 2401 ATTAAGTTTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
DB 2401 ATTAAGTTTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 ATCTTGTTCATCTCTTCAATTTTAAAGGAGGATCTGGTGGCCTGAGAAAGCAAACTTACGTC 2520
DB 2461 ATCTTGTTCATCTCTTCAATTTTAAAGGAGGATCTGGTGGCCTGAGAAAGCAAACTTACGTC 2520
QY 2521 TGGACAGATCAAAACACCCCAAGCTTGTGATGCTCAATTAATCTGTGATGATGTCGAAAAACAG 2580

1441 GCATTTGGAGTGTGAGAAATGCTATAAAGTTTCTAATCAATGGAATAAAGTGACAAC 1500
1489 GCATTTGGAGTGTGAGAAATGCTATAAAGTTTCTAATCAATGGAATAAAGTGACAAC 1548
1501 AGCAGTATGAGACCACTTCATAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
1549 AGCAGTATGAGACCACTTCATAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG 1620
1609 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG 1668
1621 TGTTCACCTTCCCGAGCCCTTCATAACCTGTGAAACACCTGTGCTTGAATGCTGATC 1680
1669 TGTTCACCTTCCCGAGCCCTTCATAACCTGTGAAACACCTGTGCTTGAATGCTGATC 1728
1681 AGAATTTGGAGTGTGGAACCATAGACATCTGGCACTTCTGTAATGCTTTGGTGAATCA 1740
1729 AGAATTTGGAGTGTGGAACCATAGACATCTGGCACTTCTGTAATGCTTTGGTGAATCA 1788
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAATACTGTTAATGGGGTCATCGCA 1800
1789 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAATACTGTTAATGGGGTCATCGCA 1848
1801 GCAGTGAACATGCTACGGAGTCTCCAGTGGCGTCTGGCTGGTGTGATGCGTTCACT 1860
1849 GCAGTGAACATGCTACGGAGTCTCCAGTGGCGTCTGGCTGGTGTGATGCGTTCACT 1908
1861 TTTGGCAGCTTTGACGACATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCATT 1920
1909 TTTGGCAGCTTTGACGACATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCATT 1968
1921 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
1969 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2028
1981 GAGCGTGGGTCTCTGTGAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
2029 GAGCGTGGGTCTCTGTGAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2088
2041 CTGAAAGTAAATCAATTTTGTCTGTCGCTGCTGGCTTGAACATGGCCGAGTCCCTG 2100
2089 CTGAAAGTAAATCAATTTTGTCTGTCGCTGCTGGCTTGAACATGGCCGAGTCCCTG 2148
2101 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGAGCCC 2160
2149 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGAGCCC 2208
2161 AGCACCATGGCTTACATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
2209 AGCACCATGGCTTACATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2268
2221 ACATTTGCTTACACCAAGCTTACTGCAATTTGGACAAGGAGACCTGGAGAAATTTGG 2280
2269 ACCATTTGCTTACACCAAGCTTACTGCAATTTGGACAAGGAGACCTGGAGAAATTTGG 2328
2281 GACTGCTCTATGTAAGACACATTTGCTGCTTGTGCTTCTTCCAACTGCTTAACTG 2340
2329 GACTGCTCTATGTAAGACACATTTGCTGCTTGTGCTTCTTCCAACTGCTTAACTG 2388
2341 CCTGTGGCTTCTTGTGCTTCTTCTTCTTAAATAAACCCTTACATTTATCACTGCTGAAGTA 2400
2389 CCTGTGGCTTCTTGTGCTTCTTCTTCTTAAATAAACCCTTACATTTATCACTGCTGAAGTA 2448
2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2449 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2508
2461 ATCTGCTTCAATCTCACTTTAAGGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2520
2509 ATCTGCTTCAATCTCACTTTAAGGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2568

2521 TGGACAGATCAAAACACCAAGCTTGATGTCATTAATCTGATGATGTCGAAAAACAG 2580
2569 TGGACAGATCAAAACACCAAGCTTGATGTCATTAATCTGATGATGTCGAAAAACAG 2628
2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
2629 TCCTGTGACTCAACTCAAGCCCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2688
2641 CTTCCAGTTCGGTGGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
2689 CTTCCAGTTCGGTGGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2748
2701 GTGGCATTTGTCCTCATGTCCTAA 2724
2749 GTGGCATTTGTCCTCATGTCCTAA 2772

RESULT 9
ADR67869
ID ADR67869 standard; cDNA; 2973 BP.
XX ADR67869;
XX AC
XX DT 18-NOV-2004 (first entry)
XX Human HG38 coding sequence.
XX ss; gene; human ; G protein-coupled receptor ; GPCR; HG38; colon ; lung ;
XX cancer.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX CDS 18..2741
XX FT /*tag= a
XX FT /product= "Human HG38"
XX WO2004074436-A2.
XX
XX PD 02-SEP-2004.
XX
XX PF 11-FEB-2004; 2004WO-US004060.
XX
XX PR 19-FEB-2003; 2003US-0448959P.
XX (INCY-) INCYTE CORP.
XX PA
XX PI Lasek AW;
XX
XX WPI; 2004-652946/63.
XX P-PSDB; ADR67868.
XX
XX PT Detecting colon or lung cancer, by determining amount of protein in
XX sample, comparing amount of protein to standard, and differential
XX expression of protein in sample indicates colon or lung cancer.
XX
XX PS Claim 1; SEQ ID NO 2; 79pp; English.
XX
XX CC This sequence encodes the human G protein-coupled receptor (GPCR) known
XX as HG38. The HG38 protein and corresponding nucleic acid, may be used in
XX the method of the invention for detecting colon or lung cancer. The first
XX method involves performing an assay to determine the amount of HG38 in a
XX sample of colon or lung tissue, and comparing the amount of protein to
XX standard, thus detecting expression of protein in sample, where
XX differential expression of protein in sample when compared with the
XX standard is diagnostic of colon or lung cancer. The second method
XX involves hybridizing a composition comprising the HG38 coding sequence,
XX or its complement, and a labelling moiety, to nucleic acids of a sample
XX of colon or lung tissue under conditions to form at least one
XX hybridization complex, detecting hybridization complex formation, and
XX comparing complex formation to a standard, where the comparison reflects
XX differential expression of the polynucleotide in the sample relative to
XX the standard and is diagnostic of a colon or lung cancer. This first

QY 241 CTGCTCCGAATCCCTCGCCAGTCTCCGCTTCCTGGAGGATTAAGTCTTGGCGGAAC 300
DB 441 CTGCTCCGAATCCCTCGCCAGTCTCCGCTTCCTGGAGGATTAAGTCTTGGCGGAAC 500
QY 301 GCTCTGACATACATTTCCCAAGGAGCATCTACTGSCCTTTTACAGTCTTAAAGTTCTTATG 360
DB 501 GCTCTGACATACATTTCCCAAGGAGCATCTACTGSCCTTTTACAGTCTTAAAGTTCTTATG 560
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
DB 561 CTGCAGAAATAATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 620
QY 421 CAATCCCTCGCTCTGGATGCTTAACACATACAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
DB 621 CAATCCCTCGCTCTGGATGCTTAACACATACAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 680
QY 481 CTGCATTCCTCGAGGACCTGTGGCTGGATGACATGGTTAAACAGAAATCCCGTCCAG 540
DB 681 CTGCATTCCTCGAGGACCTGTGGCTGGATGACATGGTTAAACAGAAATCCCGTCCAG 740
QY 541 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATATACACAC 600
DB 741 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATATACACAC 800
QY 601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCGATACAAAT 660
DB 801 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCGATACAAAT 860
QY 661 AGAATCCACTCCCTCGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 720
DB 861 AGAATCCACTCCCTCGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 920
QY 721 TTAATATTACAAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
DB 921 TTAATATTACAAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 980
QY 781 GAACTAGGATTTTATAGCAATTAATCAGCTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 981 GAACTAGGATTTTATAGCAATTAATCAGCTGATACCTGAGAAAGCAATTTGTAGGCAAC 1040
QY 841 CCTCTCTTTATTAACAATATTTCTATGACATCCCATCCCAATTTGTTGGGAGATCTGCT 900
DB 1041 CCTCTCTTTATTAACAATATTTCTATGACATCCCATCCCAATTTGTTGGGAGATCTGCT 1100
QY 901 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCAAAATAACTGAA 960
DB 1101 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCAAAATAACTGAA 1160
QY 961 TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
DB 1161 TTTCTGATTTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
QY 1021 TCATCTCTTCCCTCAAACTGCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGCT 1080
DB 1221 TCATCTCTTCCCTCAAACTGCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGCT 1280
QY 1081 TACAACCTATTAGAAGATTTACCAGATTTTTCAGTCTGCAAAAGCTTCCAGAAATTTGAC 1140
DB 1281 TACAACCTATTAGAAGATTTACCAGATTTTTCAGTCTGCAAAAGCTTCCAGAAATTTGAC 1340
QY 1141 CTAAGACATAATGAATCTACGAATTAAGTTGACATTTTCCAGCAGTTGCTTAGCCTC 1200
DB 1341 CTAAGACATAATGAATCTACGAATTAAGTTGACATTTTCCAGCAGTTGCTTAGCCTC 1400
QY 1201 CGATCGCTGAATTTGGCTTTGGAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
DB 1401 CGATCGCTGAATTTGGCTTTGGAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1460
QY 1261 TTGCCATCCCTAATAAGCTGAGCCTATCGTCCAACTCCTGTGCTTTTCTTAATTA 1320
DB 1461 TTGCCATCCCTAATAAGCTGAGCCTATCGTCCAACTCCTGTGCTTTTCTTAATTA 1520

QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
DB 1521 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1580
QY 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440
DB 1581 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1640
QY 1441 GCATTTGAGTGTGTGAGAAATGCTTAAATTAAGTATTAATCAATGGAATAAAGTGCACAA 1500
DB 1641 GCATTTGAGTGTGTGAGAAATGCTTAAATTAAGTATTAATCAATGGAATAAAGTGCACAA 1700
QY 1501 AGCAGTATGAGACGACTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
DB 1701 AGCAGTATGAGACGACTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1760
QY 1561 GACCTTGAAGATTTTCTGCTTGTACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB 1761 GACCTTGAAGATTTTCTGCTTGTACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1820
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGGCTGATC 1680
DB 1821 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGGCTGATC 1880
QY 1681 AGAATTTGAGTGTGAGACCATAGCAGTTCGGCACTTACTTGAATGCTTTGGTGAATTC 1740
DB 1881 AGAATTTGAGTGTGAGACCATAGCAGTTCGGCACTTACTTGAATGCTTTGGTGAATTC 1940
QY 1741 ACAGTTTTTACAGTCCCTCTGTACTTATTTCCCAATTTAACTGTTAAATTTGGGGTCAATCG 1800
DB 1941 ACAGTTTTTACAGTCCCTCTGTACTTATTTCCCAATTTAACTGTTAAATTTGGGGTCAATCG 2000
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGATGGCTTCACT 1860
DB 2001 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGATGGCTTCACT 2060
QY 1861 TTTGGAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAT 1920
DB 2061 TTTGGAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAT 2120
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATTCATCTGTTTTTCTGCTTCTCTGGAGCCCTG 1980
DB 2121 GGTTTTTTGTCCATTTTGTCTTCAAGATTCATCTGTTTTTCTGCTTCTCTGGAGCCCTG 2180
QY 1981 GAGCTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 2181 GAGCTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2240
QY 2041 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGCTTGAACCATGGCCGCGAGTTCCCTG 2100
DB 2241 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGCTGCTTGAACCATGGCCGCGAGTTCCCTG 2300
QY 2101 CTGGGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2160
DB 2301 CTGGGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2360
QY 2161 AGCACCATTGGCTACATGGTCTCATCTGCTTCAATTTCCCTTGTCTTCTCATGATG 2220
DB 2361 AGCACCATTGGCTACATGGTCTCATCTGCTTCAATTTCCCTTGTCTTCTCATGATG 2420
QY 2221 ACCATTGCTTACCAACAGCTCTACTGCAATTTGGAACAGGGAGACCTTGGAGAAATTTGG 2280
DB 2421 ACCATTGCTTACCAACAGCTCTACTGCAATTTGGAACAGGGAGACCTTGGAGAAATTTGG 2480
QY 2281 GACTGCTCTATGGTAAAAACATTTGCCCTGTGCTCTTCAACCACTGATCTTAACTGC 2340
DB 2481 GACTGCTCTATGGTAAAAACATTTGCCCTGTGCTCTTCAACCACTGATCTTAACTGC 2540
QY 2341 CTTGTGGCTTCTTGTGCTCTCTCTTAAATTAACCTTACATTTTATCAGTCTCTCAAGTA 2400
DB 2541 CTTGTGGCTTCTTGTGCTCTCTCTTAAATTAACCTTACATTTTATCAGTCTCTCAAGTA 2600
QY 2401 ATTAAGTTTATCTCTTCTGCTGGTAGTCCCACTTCTCTGATGCTCTCAATCCCTTCTCTAC 2460

QY 601 ATACAGACTATGCGCTTTGGAAAACTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
DB 801 ATACAGACTATGCGCTTTGGAAAACTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 860
QY 661 AGAATCCACTCCCTCGGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 861 AGAATCCACTCCCTCGGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 920
QY 721 TTAATAATACAATTAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
DB 921 TTAATAATACAATTAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 980
QY 781 GAACTAGAGATTTATAGCAACAAATATACAGTCTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 981 GAACTAGAGATTTATAGCAACAAATATACAGTCTGATACCTGAGAAAGCAATTTGTAGGCAAC 1040
QY 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCAATTTGTTGGGAGATCTGCT 900
DB 1041 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCAATTTGTTGGGAGATCTGCT 1100
QY 901 TTTCAACATTTAACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 960
DB 1101 TTTCAACATTTAACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCAAAATAACTGAA 1160
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
DB 1161 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
QY 1021 TCATCTCTTCCCTCAAACTGCTGCAATCACTGATTTACCTAATCTCCAAGTCTGATCTGCT 1080
DB 1221 TCATCTCTTCCCTCAAACTGCTGCAATCACTGATTTACCTAATCTCCAAGTCTGATCTGCT 1280
QY 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC 1140
DB 1281 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC 1340
QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTGCTGACCTC 1200
DB 1341 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTGCTGACCTC 1400
QY 1201 CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
DB 1401 CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1460
QY 1261 TTGCCATCCCTAATAAAGCTGGAACCTATGCTCCAACTCTGCTGCTTTTCCCTATAACT 1320
DB 1461 TTGCCATCCCTAATAAAGCTGGAACCTATGCTCCAACTCTGCTGCTTTTCCCTATAACT 1520
QY 1321 GGGTTACATGGTTTAACTCACTTAATAATTAACAGGAATCACTGCTTACAGAGCTTGATA 1380
DB 1521 GGGTTACATGGTTTAACTCACTTAATAATTAACAGGAATCACTGCTTACAGAGCTTGATA 1580
QY 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT 1440
DB 1581 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT 1640
QY 1441 GCATTTGGAGTGTGTGAGAAATCCCTATAGAGATTTCTAATCAATGAATAAAGGTGACAAAC 1500
DB 1641 GCATTTGGAGTGTGTGAGAAATCCCTATAGAGATTTCTAATCAATGAATAAAGGTGACAAAC 1700
QY 1501 AGCAGTATGGACGACTTTCATTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGT 1560
DB 1701 AGCAGTATGGACGACTTTCATTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGT 1760
QY 1561 GACCTTGAAGATTTCTGCTGCTGACTTTGAGGAAGACCTGAAAAGCCCTTCATTCAGTGCAG 1620
DB 1761 GACCTTGAAGATTTCTGCTGCTGACTTTGAGGAAGACCTGAAAAGCCCTTCATTCAGTGCAG 1820
QY 1621 TGTTCACTTCCCGAGGCCCTTTCAAACTCTGTGAAACCACTGCTTGTAGTGGCTGGCTGATC 1680
DB 1821 TGTTCACTTCCCGAGGCCCTTTCAAACTCTGTGAAACCACTGCTTGTAGTGGCTGGCTGATC 1880

QY 1681 AGAATTCGAGTGTGGAACCATAGACAGTTCTGSCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
DB 1881 AGAATTCGAGTGTGGAACCATAGACAGTTCTGSCACTTACTTGTAAATGCTTTGGTGACTTCA 1940
QY 1741 ACAGTTTTTCAGATCCCTCTGTGTACATTTCCCCCACTTAAACTGTTAAATTTGGGGTCAATCGCA 1800
DB 1941 ACAGTTTTTCAGATCCCTCTGTGTACATTTCCCCCACTTAAACTGTTAAATTTGGGGTCAATCGCA 2000
QY 1801 GCAGTGAACATGCTCAAGGGAGTCTCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 2001 GCAGTGAACATGCTCAAGGGAGTCTCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2060
QY 1861 TTTGCGAGCTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 2061 TTTGCGAGCTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 2121 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 2180
QY 1981 GAGCGTGGTCTCTGCTGAAATATTTGCAAAAATTTGAAACGAAAGCTCCATTTTCTTAGC 2040
DB 2181 GAGCGTGGTCTCTGCTGAAATATTTGCAAAAATTTGAAACGAAAGCTCCATTTTCTTAGC 2240
QY 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2241 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300
QY 2101 CTGGGTGGCAGCAAGTATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2301 CTGGGTGGCAGCAAGTATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2360
QY 2161 AGCAGCATGGCTACATGCTGCTCTCATCTGCTGCTCAATTTCCCTTTCCTTCTCATGATG 2220
DB 2361 AGCAGCATGGCTACATGCTGCTCTCATCTGCTGCTCAATTTCCCTTTCCTTCTCATGATG 2420
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTTGACAAGGAGAGACTGGAGAAATATTTGG 2280
DB 2421 ACCATTGCTACACCAAGCTCTACTGCAATTTTGACAAGGAGAGACTGGAGAAATATTTGG 2480
QY 2281 GACTGCTCTATGTTGTAACACATTTGCCCTGTTCTTTACCAACTGCAATCTCTAAACTGTC 2340
DB 2481 GACTGCTCTATGTTGTAACACATTTGCCCTGTTCTTTACCAACTGCAATCTCTAAACTGTC 2540
QY 2341 CCTGCTGCTTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
DB 2541 CCTGCTGCTTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2600
QY 2401 ATTAAGTTTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
DB 2601 ATTAAGTTTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2660
QY 2461 ATCTGTTTCAATCTCTCACTTTAAGAGGATCTGCTGAGCCTTGAGAAAGCAAACTTACGTC 2520
DB 2661 ATCTGTTTCAATCTCTCACTTTAAGAGGATCTGCTGAGCCTTGAGAAAGCAAACTTACGTC 2720
QY 2521 TGGACAAGATCAAAACACCAAGCTTGTATGCTCAATTAACCTCTGATGCTGCAAAAACAG 2580
DB 2721 TGGACAAGATCAAAACACCAAGCTTGTATGCTCAATTAACCTCTGATGCTGCAAAAACAG 2780
QY 2581 TCCTGTGACTCAACTCAAGCCTTGTAACTTACCAGCTCCAGCATCACTTATGACCTG 2640
DB 2781 TCCTGTGACTCAACTCAAGCCTTGTAACTTACCAGCTCCAGCATCACTTATGACCTG 2840
QY 2641 CCTCCAGTTCGCTGCCATCAAGCTTATCCAGTGTATCCAGTGTAGAGCTGCTGCTTCTCTCT 2700
DB 2841 CCTCCAGTTCGCTGCCATCAAGCTTATCCAGTGTATCCAGTGTAGAGCTGCTGCTTCTCTCT 2900
QY 2701 GTGGCAATTTGCTCCATGCTCTAA 2724
DB 2901 GTGGCAATTTGCTCCATGCTCTAA 2924

[illegible]

Qy	1021	TCATCTCTTCTCTGAAACCGTCTGGACATGATTAACCTTCCAAAGTGTCTAGATCTGTCT	1081
Db	1282	TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTAACTTCCAAAGTGTCTAGATCTGTCT	1341
Qy	1081	TACAACCTATTAGAGAATTTTACCCAGTTTTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC	1140
Db	1342	TACAACCTATTAGAGAATTTTACCCAGTTTTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC	1401
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACATTTTCCAGCAGTTTGTCTAGCCTC	1200
Db	1402	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACATTTTCCAGCAGTTTGTCTAGCCTC	1461
Qy	1201	CGATCGCTGAATTTTGGCTTGGAAACAAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT	1260
Db	1462	CGATCGCTGAATTTTGGCTTGGAAACAAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT	1521
Qy	1261	TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTGCTTTTCTTTCTATAA	1320
Db	1522	TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTGCTTTTCTTTCTATAA	1581
Qy	1321	GGGTTACATGGTTTAACTCACTTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1582	GGGTTACATGGTTTAACTCACTTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1641
Qy	1381	TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCAGTGTCTGT	1440
Db	1642	TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCAGTGTCTGT	1701
Qy	1441	GCATTTGGAGTGTGTGAGATGCTTATAGATTTCTTAATCAATGGAATAAGGTGACAACT	1500
Db	1702	GCATTTGGAGTGTGTGAGATGCTTATAGATTTCTTAATCAATGGAATAAGGTGACAACT	1761
Qy	1501	AGCAGTATGACGACCTTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1762	AGCAGTATGACGACCTTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1821
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTGAGGAGAACTGAAAGCCCTTCATTCAAGTGCAG	1620
Db	1822	GACCTTGAAGATTTCTGCTTGACTTTGAGGAGAACTGAAAGCCCTTCATTCAAGTGCAG	1881
Qy	1621	TGTTCACTTCCCGAGGCCCTTCAAACTCTGTGAAACCTGCTGATGCTGGCTGGCTGATC	1680
Db	1882	TGTTCACTTCCCGAGGCCCTTCAAACTCTGTGAAACCTGCTGATGCTGGCTGGCTGATC	1941
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATTCA	1740
Db	1942	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATTCA	2001
Qy	1741	ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAACCTGTTAAATGGGGTCATCGCA	1800
Db	2002	ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAACCTGTTAAATGGGGTCATCGCA	2061
Qy	1801	GCAGTGAACATGCTCAGGGAGTCTCAGTGGCCGTGCTGGCTGGTGTGAGATGGTTCACT	1860
Db	2062	GCAGTGAACATGCTCAGGGAGTCTCAGTGGCCGTGCTGGCTGGTGTGAGATGGTTCACT	2121
Qy	1861	TTTGGCAGCTTTTGCACGACATGTGCTGTGTGGAGAAATGGGGTGTGGTTCATGTCATT	1920
Db	2122	TTTGGCAGCTTTTGCACGACATGTGCTGTGTGGAGAAATGGGGTGTGGTTCATGTCATT	2181
Qy	1921	GGTTTTTGTGTCATTTTGTCTCAGAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	2182	GGTTTTTGTGTCATTTTGTCTCAGAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	2241
Qy	1981	GAGCGTGGGTTCTCTGTGAATAATTTCTGMAAATTTTGAACGAAAGCTCAATTTTCTAGC	2040
Db	2242	GAGCGTGGGTTCTCTGTGAATAATTTCTGMAAATTTTGAACGAAAGCTCAATTTTCTAGC	2301
Qy	2041	CTGAAAGTAAATCAATTTTGTCTGTGCCCTGTGSCCTTGACCAATGCGCAGTTTCCCTTG	2100
Db	2302	CTGAAAGTAAATCAATTTTGTCTGTGCCCTGTGSCCTTGACCAATGCGCAGTTTCCCTTG	2361
Qy	2101	CTGGGTGGCAGCAAGTATGSCGCTTCCCTCTCTGCTGCTTTTGTGCTTTTGGGAGCCCT	2160

Db 2362 CTGGGTGGAGCAAGATATGGCCCTCCCTCTCTCCCTGGCTTGGCCCTTTGGGGAGGCC 2421
Qy 2161 AGCACCATGGGCTACATGTCGCTCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2220
Db 2422 AGCACCATGGGCTACATGTCGCTCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2481
Qy 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATTTGG 2280
Db 2482 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATTTGG 2541
Qy 2281 GACTGCTCTATGTAATAACACATTCGCTTGTCTTTCACCAACTGCATCTTAACTGC 2340
Db 2542 GACTGCTCTATGTAATAACACATTCGCTTGTCTTTCACCAACTGCATCTTAACTGC 2601
Qy 2341 CCGTGGGCTTTCTGTCTCTCTCTCTCTCTTAAATAACCTTTACATTTATCAGTCTCGAAGTA 2400
Db 2602 CCGTGGGCTTTCTGTCTCTCTCTCTCTCTTAAATAACCTTTACATTTATCAGTCTCGAAGTA 2661
Qy 2401 ATTAAGTTTATCTTCTGTGTGTAGTCCGACTTCTGCTGATGTCATCCCTTCTCTAC 2460
Db 2662 ATTAAGTTTATCTTCTGTGTGTAGTCCGACTTCTGCTGATGTCATCCCTTCTCTAC 2721
Qy 2461 ATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGGTGAGCTTGAGAAAGCAACCTTACGTC 2520
Db 2722 ATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGGTGAGCTTGAGAAAGCAACCTTACGTC 2781
Qy 2521 TGGACAAGATCAAAACACCAAGCTTGATGTCATTAATTAATCTCTGATGATGTGCAAAAAACAG 2580
Db 2782 TGGACAAGATCAAAACACCAAGCTTGATGTCATTAATTAATCTCTGATGATGTGCAAAAAACAG 2841
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGTTAAGCTTTTACAGTCCAGCTCAGCATCACTTATGACCTG 2640
Db 2842 TCCTGTGACTCAACTCAAGCCTTGTTAAGCTTTTACAGTCCAGCTCAGCATCACTTATGACCTG 2901
Qy 2641 CTCTCCAGTTCCGTGCGATCACAGCTTATCCAGTGACTGAGAGCTGCGATCTTTCTCTCT 2700
Db 2902 CTCTCCAGTTCCGTGCGATCACAGCTTATCCAGTGACTGAGAGCTGCGATCTTTCTCTCT 2961
Qy 2701 GTGGCATTTGTCCCATGTCTCTAA 2724
Db 2962 GTGGCATTTGTCCCATGTCTCTAA 2985

RESULT 14

ADF70582
ID ADF70582 standard; DNA; 3438 BP.

AC ADF70582;
XX

DT 12-FEB-2004 (first entry)
XX

DE Orphan receptor ligand-related human protein gene SeqID205.
XX

XX ligand; orphan receptor protein; fusion protein; fluorescent protein;
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
XX

OS Homo sapiens.
XX

XX WO2003071272-A1.
XX

XX 28-AUG-2003.
XX

XX 21-FEB-2003; 2003WO-JP001901.
XX

XX 22-FEB-2002; 2002JP-00045728.
XX

XX 23-JUL-2002; 2002JP-00213949.
XX

XX 11-OCT-2002; 2002JP-00298237.
XX

XX (TAKE) TAKEDA CHEM IND LTD.
XX

XX Hinuma S, Fujii R, Ogi K, Komateu H, Kawamata Y, Hoboya M;
PI

XX WPI; 2003-697654/66.
DR P-PSDB; ADF70480.
XX Transformation of cells with a fusion protein of an orphan receptor
PT protein with a fluorescent protein useful for identification of ligands
PT to the orphan receptor.
XX Example 4; SEQ ID NO 205; 594pp; Japanese.
PS This invention relates to a novel method of identifying ligands to an
CC orphan receptor protein which comprises transforming cells with DNA
CC encoding a fusion protein of the orphan receptor with a fluorescent
CC protein, so that the fusion protein is expressed in the cells (or cell
CC membranes isolated from them) and contacting the cells with the potential
CC ligand to be tested. A suitable fluorescent protein for incorporation in
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
CC identification of ligands binding to an orphan receptor protein.
XX Sequence 3438 BP; 894 A; 881 C; 707 G; 956 T; 0 U; 0 Other;
SQ

Query Match 98.0%; Score 2670; DB 10; Length 3438;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2720; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGACACCTCCCGCTCGGTGCTCTCTGCTCTTGGCTGTGCTGCAGCTGGCGACC 60
Db 1 ATGGACACCTCCCGCTCGGTGCTCTCTGCTCTTGGCTGTGCTGCAGCTGGCGACC 60
Qy 61 GGGGCGAGCTCTCCCGCTCGGTGCTCTGCTGAGGGGCTGCCCCACACACTGTCATTGC 120
Db 61 GGGGCGAGCTCTCCCGCTCGGTGCTCTGCTGAGGGGCTGCCCCACACACTGTCATTGC 120
Qy 121 GAGCCGACGCGCAGGATGTTGCTCAGGGTGACATGCTCCGACCTGGGGCTCTCGAGAGCTG 180
Db 121 GAGCCGACGCGCAGGATGTTGCTCAGGGTGACATGCTCCGACCTGGGGCTCTCGAGAGCTG 180
Qy 181 CTTTCCAACTCAGGGTCTTCACTCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 181 CTTTCCAACTCAGGGTCTTCACTCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTTACGTTCTTGGCGGAAC 300
Db 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTTACGTTCTTGGCGGAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Qy 361 CTGCAGATATACAGCTTAAGACAGCTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 CTGCAGATATACAGCTTAAGACAGCTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Qy 421 CAATCCCTGCTCTGGATGCTAAACACATCAGTATGTGCCCCCAAGCTGTTTCAAGTGGC 480
Db 421 CAATCCCTGCTCTGGATGCTAAACACATCAGTATGTGCCCCCAAGCTGTTTCAAGTGGC 480
Qy 481 CTGCATTTCCCTGAGGCACCTGTGGTGATGACAATGCGTTAAAGAAATCCCGTCCAG 540
Db 481 CTGCATTTCCCTGAGGCACCTGTGGTGATGACAATGCGTTAAAGAAATCCCGTCCAG 540
Qy 541 GCTTTTAGAAGTTTATCGGCATTCGACCTGACCTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTAGAAGTTTATCGGCATTCGACCTGACCTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720

Db 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTTCTATAACT 1320
Qy
1321 GGGTTACATGGTTTAACTCACTTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
Db
1321 GGGTTACATGGTTTAACTCACTTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
Qy
1381 TCATCTGAAACTTTCCAGNACTCAAGGTATAGAAATGCGTTATGCTTACCGAGTCTGT 1440
Db
1381 TCATCTGAAACTTTCCAGNACTCAAGGTATAGAAATGCGTTATGCTTACCGAGTCTGT 1440
Qy
1441 GCATTTGGAGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Db
1441 GCATTTGGAGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Qy
1501 AGCAGTATGGAACACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db
1501 AGCAGTATGGAACACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy
1561 GACCTTGAAATTTCCCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db
1561 GACCTTGAAATTTCCCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Qy
1621 TGTTCACCTTTCCCGAGGCCCTTCAAAACCTGTGAAACACCTGCTTCATGCTGCTGATC 1680
Db
1621 TGTTCACCTTTCCCGAGGCCCTTCAAAACCTGTGAAACACCTGCTTCATGCTGCTGATC 1680
Qy
1681 AGAATTTGGAGTGTGGAACCATAGAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db
1681 AGAATTTGGAGTGTGGAACCATAGAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Qy
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Db
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Qy
1801 GCAGTGAAACATGCTCAGGGAGTCTCCAGTGGCGTCTGCTGCTGTGATGCTGTTCACT 1860
Db
1801 GCAGTGAAACATGCTCAGGGAGTCTCCAGTGGCGTCTGCTGCTGTGATGCTGTTCACT 1860
Qy
1861 TTTGGCAGCTTTGACGACATGGTGTGCTGGTGGAGAATGGGGTGGTTGCCATGTCAAT 1920
Db
1861 TTTGGCAGCTTTGACGACATGGTGTGCTGGTGGAGAATGGGGTGGTTGCCATGTCAAT 1920
Qy
1921 GGTTTTGTGCTCAATTTTGTGCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db
1921 GGTTTTGTGCTCAATTTTGTGCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
Qy
1981 GAGCGTGGTCTCTCTGGAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db
1981 GAGCGTGGTCTCTCTGGAATAATCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Qy
2041 CTGAAAGTAATCATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db
2041 CTGAAAGTAATCATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy
2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db
2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Qy
2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db
2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Qy
2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTGGAGAATAATTTGG 2280
Db
2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTGGAGAATAATTTGG 2280
Qy
2281 GACTGCTCTATGTTAAACACATGCGCTGTTGCTCTTCAACCACTGCATCTCTAAACTGC 2340
Db
2281 GACTGCTCTATGTTAAACACATGCGCTGTTGCTCTTCAACCACTGCATCTCTAAACTGC 2340
Qy
2341 CCTGCGCTTCTGCTGCT 2400
Db
2341 CCTGCGCTTCTGCTGCT

Qy 2401 ATTAAGTTTATCCTTTCTGTTGTAAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Db
2401 ATTAAGTTTATCCTTTCTGTTGTAAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Qy 2461 ATCTTGTTCATCCTCACTTTTAAAGGAGATCTGTTGAGCTGAGAAAGCAAACTTACGTC 2520
Db
2461 ATCTTGTTCATCCTCACTTTTAAAGGAGATCTGTTGAGCTGAGAAAGCAAACTTACGTC 2520
Qy 2521 TGGACAAGATCAAAACACCCAAAGCTTGTGATGCTCAATTAATCTGATGATGTCGAAAAACAG 2580
Db
2521 TGGACAAGATCAAAACACCCAAAGCTTGTGATGCTCAATTAATCTGATGATGTCGAAAAACAG 2580
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTAACAGCTCCAGCATCACTTATGACCTG 2640
Db
2581 TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTAACAGCTCCAGCATCACTTATGACCTG 2640
Qy 2641 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTCT 2700
Db
2641 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTCT 2700
Qy 2701 GTGGCATTTGTCCCATGTCTCTAA 2724
Db
2701 GTGGCATTTGTCCCATGTCTCTAA 2724

Search completed: July 12, 2005, 21:29:18
Job time : 1423 secs

This Page Blank (uspto)

||||| 541 GCCTTTAGAAAGTTTATCGGCAATTGCAAGCATGACCTTGGCCCTGAAACAAAATACACCAC 600
Db
QY 601 ATACCGAGATATGCTTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
Db
QY 601 ATACCGAGATATGCTTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
Db
QY 661 AGAATCGATCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGATTTAGAT 720
Db
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGATTTAGAT 720
Db
QY 721 TTAATTTACAAATAACCTTGTAGTAATCCCACTGCAATTAGGACACTCTCCAACTTTAA 780
Db
QY 721 TTAATTTACAAATAACCTTGTAGTAATCCCACTGCAATTAGGACACTCTCCAACTTTAA 780
Db
QY 781 GAACTAGGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db
QY 841 CCTTCTCTTATTAACAATAATTTCTATGACAATCCCACTGCAATTTGTAGGCAAC 840
Db
QY 841 CCTTCTCTTATTAACAATAATTTCTATGACAATCCCACTGCAATTTGTAGGCAAC 840
Db
QY 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
Db
QY 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
Db
QY 961 TTTCTGTGATTTAACTGGAACCTGAAACCTGGAGAGTCTGACTTTAACTGAGACACAGATC 1020
Db
QY 961 TTTCTGTGATTTAACTGGAACCTGAAACCTGGAGAGTCTGACTTTAACTGAGACACAGATC 1020
Db
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAACTCCCAAGTCTAGATCTGTCT 1080
Db
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAACTCCCAAGTCTAGATCTGTCT 1080
Db
QY 1081 TACAACCTTATAGAAAGTTTACCAGTTTTTCCAGTCTGCAAAAGCTTCAGAAAATTGAC 1140
Db
QY 1081 TACAACCTTATAGAAAGTTTACCAGTTTTTCCAGTCTGCAAAAGCTTCAGAAAATTGAC 1140
Db
QY 1141 CTAAGACATAAATAAGTACGAAATTTAAAGTTTGAACACTTTCCAGCAGTTTGTAGCCCTC 1200
Db
QY 1141 CTAAGACATAAATAAGTACGAAATTTAAAGTTTGAACACTTTCCAGCAGTTTGTAGCCCTC 1200
Db
QY 1201 CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTCAACCCCAATGCAATTTCCACT 1260
Db
QY 1261 TTTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db
QY 1321 GGGTTACATGGTTTAACTCACTTTAAATTTAAACAGGAAATCATGSCCTTTACAGAGCTTGATA 1380
Db
QY 1321 GGGTTACATGGTTTAACTCACTTTAAATTTAAACAGGAAATCATGSCCTTTACAGAGCTTGATA 1380
Db
QY 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTTACCAGTCTGT 1440
Db
QY 1441 GCATTTGGAGTGTGAGAAATGCTTATAGATTTCTAATCAATGGAATTAAGGTGACAAAC 1500
Db
QY 1501 AGCAGTATGACCACTTTCATAGAAAGATGCTTGAATGTTTTCCAGCTCAAGATGAACCT 1560
Db
QY 1501 AGCAGTATGACCACTTTCATAGAAAGATGCTTGAATGTTTTCCAGCTCAAGATGAACCT 1560
Db
QY 1561 GACCTTTGAGATTTCTGCTTGAATTTGAGGAGACCTTGAAAGCCCTTCATTCAGTGCAG 1620
Db
QY 1561 GACCTTTGAGATTTCTGCTTGAATTTGAGGAGACCTTGAAAGCCCTTCATTCAGTGCAG 1620
Db
QY 1621 TGTTCACTTCCCGGCCCTTTCAAAACCTGTGAAACACTGCTTGAATGGCTGGCTGATC 1680

1621 TGTTCACTTCCCGGCCCTTTCAAAACCTGTGAAACACTGCTTGAATGGCTGGCTGATC 1680
QY
1681 AGAATTTGAGTGTGGACCATAGCAGTTCTGGCACTTTACTTTGTAATGCTTTTGGTGACTTCA 1740
Db
1681 AGAATTTGAGTGTGGACCATAGCAGTTCTGGCACTTTACTTTGTAATGCTTTTGGTGACTTCA 1740
QY
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGGTCACTGCA 1800
Db
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGGTCACTGCA 1800
QY
1801 CGAGTGAACATCTCACGGGAGTCTCCAGTGCCTGCTGGTGGTGTGATGATGCTTCACT 1860
Db
1801 CGAGTGAACATCTCACGGGAGTCTCCAGTGCCTGCTGGTGGTGTGATGATGCTTCACT 1860
QY
1861 TTTGGCAGCTTTTGCACGACATGCTGCTGGTGGGAAATGGGGTTGGTTGCCATGTCAAT 1920
Db
1861 TTTGGCAGCTTTTGCACGACATGCTGCTGGTGGGAAATGGGGTTGGTTGCCATGTCAAT 1920
QY
1921 GGTTTTTTGTCCATTTTGTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGGAGGCCCTG 1980
Db
1921 GGTTTTTTGTCCATTTTGTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGGAGGCCCTG 1980
QY
1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTTCAATTTTCTAGC 2040
Db
1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTTCAATTTTCTAGC 2040
QY
2041 CTGAAAGTAAATCATTTTGTCTGTGCTGCTGCTGGCTTGGCATGCGCCAGTTCCTCCCTG 2100
Db
2041 CTGAAAGTAAATCATTTTGTCTGTGCTGCTGCTGGCTTGGCATGCGCCAGTTCCTCCCTG 2100
QY
2101 CTGGTGGCAGCAAGATGAGGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGGAGCCC 2160
Db
2101 CTGGTGGCAGCAAGATGAGGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGGAGCCC 2160
QY
2161 AGCACATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db
2161 AGCACATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
QY
2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATTTGG 2280
Db
2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATTTGG 2280
QY
2281 GACTGCTATATGGTAAACAACATGCTGCTGCTTCTTCAACCACTGCAATTCCTAACTGC 2340
Db
2281 GACTGCTATATGGTAAACAACATGCTGCTGCTTCTTCAACCACTGCAATTCCTAACTGC 2340
QY
2341 CCTGTGGCTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db
2341 CCTGTGGCTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
QY
2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db
2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
QY
2461 ATCTTGTTCATTCCTCACTTTTAAAGGAGATCTGGTGGCTTGAAGGAGGAGGAGGAGGAGG 2520
Db
2461 ATCTTGTTCATTCCTCACTTTTAAAGGAGATCTGGTGGCTTGAAGGAGGAGGAGGAGGAGG 2520
QY
2521 TGGACAAGATCAAAACACCAAGCTTGTGATGCTTCAATTAATCTGATGATGCTGCAAAACAG 2580
Db
2521 TGGACAAGATCAAAACACCAAGCTTGTGATGCTTCAATTAATCTGATGATGCTGCAAAACAG 2580
QY
2581 TCCTGTGACTCAACTCAAGGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db
2581 TCCTGTGACTCAACTCAAGGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY
2641 CCTCCAGTTCGGTGCCATCACCAGCTTATCCAGTGAATGAGAGCTGCAATCTTCTTCTTCT 2700
Db
2641 CCTCCAGTTCGGTGCCATCACCAGCTTATCCAGTGAATGAGAGCTGCAATCTTCTTCTTCT 2700
QY
2701 GTGGCAATTTGCTCCATGCTCTTAA 2724
Db
2701 GTGGCAATTTGCTCCATGCTCTTAA 2724


```

RESULT 4
US-09-495-050A-220
; Sequence 220, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495.050A
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 220
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2603450CT1
US-09-495-050A-220

Query Match      25.7%; Score 699; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1685  TTGAGGTGTGGACCATAGCAGTCTCTGGCAGCTACTCTGTAATGCTTTGGTGACTTCAACAG 1744
Db
Qy      1745  TTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAATGCAGCAG 1804
Db      61  TTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAATGCAGCAG 120
Qy      1805  TGAACATGCTCACGGAGTCTCCAGTGCCGCTGCTGGTGCGTGGTGGAATGCGTTCACATTTTG 1864
Db      121  TGAACATGCTCACGGAGTCTCCAGTGCCGCTGCTGGTGCGTGGTGGAATGCGTTCACATTTTG 180
Qy      1865  GCAGCTTGGACAGCATGCTGCTCGTGGGAGAAATGGGTTGGTTGGCATGTCTCATTTGGTT 1924
Db      181  GCAGCTTGGACAGCATGCTGCTGGTGGAATGGGTTGGTTGGCATGTCTCATTTGGTT 240
Qy      1925  TTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGC 1984
Db      241  TTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGC 300
Qy      1985  GTGGGTTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCCAATTTCTAGCCCTGA 2044
Db      301  GTGGGTTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCCAATTTCTAGCCCTGA 360
Qy      2045  AAGTAATCATTTTGTCTGTGCGCTGCTGGCTTTGACCATGGCGCAGTTCCCTCTGCTGG 2104
Db      361  AAGTAATCATTTTGTCTGTGCGCTGCTGGCTTTGACCATGGCGCAGTTCCCTCTGCTGG 420
Qy      2105  GTGGCAGCAAGTATGGGCGCTCCCTCTCTGCGCTTTTGGGAGAGCCCAAGCA 2164
Db      421  GTGGCAGCAAGTATGGGCGCTCCCTCTCTGCGCTTTTGGGAGAGCCCAAGCA 480
Qy      2165  CCATGGGGTACATGGTGTGCTCTCATCTTGCTCAATTCGCTTGTCTCTCATGATGACCA 2224
Db      481  CCATGGGGTACATGGTGTGCTCTCATCTTGCTCAATTCGCTTGTCTCTCATGATGACCA 540
Qy      2225  TTGGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATTTTGGGACT 2284
Db      541  TTGGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATTTTGGGACT 600
Qy      2285  GCTCTATGGTAAAAACAAATTTGCCCTGTGCTCTTTCACCAACTGCATCTTAACTGCCCTG 2344

```

```
; Sequence 5, Application US/08971089
; Patent No. 6376174
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; APPLICANT: Scoles, Daniel R.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING
; TITLE OF INVENTION: SCHWANNIN-BINDING-PROTEINS AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,089
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,987
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramos, Robert T.
; REGISTRATION NUMBER: 37,915
; REFERENCE/DOCKET NUMBER: P-CE 2862
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..2077
; US-08-971-089-5

Query Match 0.8%; Score 21; DB 3; Length 2665;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATTGGGCTACATGG 2179
Db 1815 CCAGCACCATTGGGCTACATGG 1835

RESULT 8
US-09-266-225D-9
; Sequence 9, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingmore, Stephen
; APPLICANT: Tchernov, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2896

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-266-225D-9

Query Match 0.8%; Score 21; DB 4; Length 2896;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATTGGGCTACATGG 2179
Db 2132 CCAGCACCATTGGGCTACATGG 2152

RESULT 9
US-09-949-016-669
; Sequence 669, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669
; LENGTH: 2906
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-669

Query Match 0.8%; Score 21; DB 4; Length 2906;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATTGGGCTACATGG 2179
Db 2147 CCAGCACCATTGGGCTACATGG 2167

RESULT 10
US-09-949-016-4412
; Sequence 4412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4412
; LENGTH: 2910
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-4412

Query Match 0.8%; Score 21; DB 4; Length 2910;
```

```
Best Local Similarity 100.0%; Pred. No. 3.6; DB 4; Length 22339;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
|||||
Db 2151 CCAGCACCATGGGCTACATGG 2171

RESULT 11
US-09-949-016-12411
; Sequence 12411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12411
; LENGTH: 22339
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22339)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12411

Query Match 0.8%; Score 21; DB 4; Length 22339;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
|||||
Db 18974 CCAGCACCATGGGCTACATGG 18994

RESULT 12
US-09-949-016-16154
; Sequence 16154, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16154
; LENGTH: 22339
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22339)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-949-016-16154
Query Match 0.8%; Score 21; DB 4; Length 22339;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
|||||
Db 18974 CCAGCACCATGGGCTACATGG 18994

RESULT 13
US-09-949-016-17248
; Sequence 17248, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17248
; LENGTH: 61198
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(61198)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17248

Query Match 0.8%; Score 21; DB 4; Length 61198;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1994 CTGTGAAATATTTCTGCAAAAT 2014
|||||
Db 35060 CTGTGAAATATTTCTGCAAAAT 35080

RESULT 14
US-09-621-976-13818/c
; Sequence 13818, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13818
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 405
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13818
```

Query Match 0.7%; Score 20; DB 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 CCTGCCCGAGTCTCCGCTTCC 274
|||||
Db 315 CCTGCCCGAGTCTCCGCTTCC 296

RESULT 15

US-09-270-767-14190
; Sequence 14190, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14190
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14190

Query Match 0.7%; Score 20; DB 4; Length 574;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 715 TTAGATTAAATTACAATTA 734
|||||
Db 415 TTAGATTAAATTACAATTA 434

Search completed: July 13, 2005, 03:13:17
Job time : 457 secs

This page Blank (uspto)

Db	241		CTGCTCCGGAATCCCCCTGCCAGTCTCCGCTCTCTGGAGGATTAAGTCTTGGCGGGAAC	300
Qy	301	GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTTACAGTCTTAAAGTTCTTTATG		360
Db	301	GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTTACAGTCTTAAAGTTCTTTATG		360
Qy	361	CTGCAGAAATAATCAGCTAAGACACGTTACCCACAGAAGCTCTGCGAAGATTTGCGAAGCCTT		420
Db	361	CTGCAGAAATAATCAGCTAAGACACGTTACCCACAGAAGCTCTGCGAAGATTTGCGAAGCCTT		420
Qy	421	CAATCCCTGGCTCTGGATGCTTAAACCACATGAGCTATGTGCCCCCAAGCTGTTTCAGTGGC		480
Db	421	CAATCCCTGGCTCTGGATGCTTAAACCACATGAGCTATGTGCCCCCAAGCTGTTTCAGTGGC		480
Qy	481	CTGCATTTCCCTGAGGCACCTGTGCTGATGACAAATGCTTTAAACAGAAATCCCGCTCCAG		540
Db	481	CTGCATTTCCCTGAGGCACCTGTGCTGATGACAAATGCTTTAAACAGAAATCCCGCTCCAG		540
Qy	541	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATAACCCAC		600
Db	541	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATAACCCAC		600
Qy	601	ATACGAGACTATGCTTTGGAAACCTCTCAGCTTTGTTAGTTCTPA CATCTCCATAACAAT		660
Db	601	ATACGAGACTATGCTTTGGAAACCTCTCAGCTTTGTTAGTTCTPA CATCTCCATAACAAT		660
Qy	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT		720
Db	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT		720
Qy	721	TTAAATTAACAATAACCTTTGATGAAATCCCCACTGCAATTAGGACACCTCTCCAACTTTAAA		780
Db	721	TTAAATTAACAATAACCTTTGATGAAATCCCCACTGCAATTAGGACACCTCTCCAACTTTAAA		780
Qy	781	GAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC		840
Db	781	GAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC		840
Qy	841	CTCTCTCTTATTACAATATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGATCTGCT		900
Db	841	CTCTCTCTTATTACAATATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGATCTGCT		900
Qy	901	TTTCAACATTTTACCTGAACTAAGAACACCTGACTCTGAAATGGTGCTCACAAATAACTGAA		960
Db	901	TTTCAACATTTTACCTGAACTAAGAACACCTGACTCTGAAATGGTGCTCACAAATAACTGAA		960
Qy	961	TTTCCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC		1020
Db	961	TTTCCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC		1020
Qy	1021	TCACTCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTGCTTAGATCTGCT		1080
Db	1021	TCACTCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTGCTTAGATCTGCT		1080
Qy	1081	TACAACCTATTAGAAGATTTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC		1140
Db	1081	TACAACCTATTAGAAGATTTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC		1140
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTGCTTAGCCTC		1200
Db	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTGCTTAGCCTC		1200
Qy	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTCCACT		1260
Db	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTCCACT		1260
Qy	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCGTCCTTTTCTCTATAACT		1320
Db	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCGTCCTTTTCTCTATAACT		1320
Qy	1321	GGGTTTACATGGTTTAACTCACTTAAAAATTTAAACAGGAAATCATGCTCTTACAGAGCTTGATA		1380

Db	1321	GGGTTA	CATGGTTTAA	CTCACTTAA	AAATTAACAGGA	AATCATG	CGCTTTA	CAGAGCTTG	GA	1381
Qy	1381	TCATCTG	AAAACTTT	CCAGAACT	CAAGGTTTATAG	AAATGCTTTATG	CTTACCAGT	GC	TGT	1440
Db	1381	TCATCTG	AAAACTTT	CCAGAACT	CAAGGTTTATAG	AAATGCTTTATG	CTTACCAGT	GC	TGT	1440
Qy	1441	GCATTTGG	AGTG	TGTGAGAA	TGCTTATAG	ATTTCTAAT	CAATGGA	ATAAAGGTG	ACAAC	1500
Db	1441	GCATTTGG	AGTG	TGTGAGAA	TGCTTATAG	ATTTCTAAT	CAATGGA	ATAAAGGTG	ACAAC	1500
Qy	1501	AGCAGTATG	ACGACCTT	CTATAGA	AGATGCTTGG	AAATGTTCTCAGG	CTCAAGAT	GAC	AGT	1560
Db	1501	AGCAGTATG	ACGACCTT	CTATAGA	AGATGCTTGG	AAATGTTCTCAGG	CTCAAGAT	GAC	AGT	1560
Qy	1561	GACCTTGA	AGATTTCTG	CTTGACTTTG	AGGAAGACCTG	AAAGCCCTTCA	TTCAAGTGC	AG		1620
Db	1561	GACCTTGA	AGATTTCTG	CTTGACTTTG	AGGAAGACCTG	AAAGCCCTTCA	TTCAAGTGC	AG		1620
Qy	1621	TGTTCACT	TTCCGAGG	CCCTTCA	AAACCTGTG	AAACCTGTG	AAACCTGTG	AAACCTGTG	AAACCTGTG	1680
Db	1621	TGTTCACT	TTCCGAGG	CCCTTCA	AAACCTGTG	AAACCTGTG	AAACCTGTG	AAACCTGTG	AAACCTGTG	1680
Qy	1681	AGAAATGG	AGTG	GGAACATAG	CAGATCTTG	GGCACTTCTG	TGTAATG	CTTTGGT	TGACTTCA	1740
Db	1681	AGAAATGG	AGTG	GGAACATAG	CAGATCTTG	GGCACTTCTG	TGTAATG	CTTTGGT	TGACTTCA	1740
Qy	1741	ACAGTTTTC	CAGATCCC	CTCTGTAC	ATTTCC	CCCAATTA	ATTTGGGG	CTCATG	CGCA	1800
Db	1741	ACAGTTTTC	CAGATCCC	CTCTGTAC	ATTTCC	CCCAATTA	ATTTGGGG	CTCATG	CGCA	1800
Qy	1801	GCAGTGA	ACATGCT	CACGG	AGTCTC	ACAGTGC	GTGCTGGT	TGGAATG	CGTTCACT	1860
Db	1801	GCAGTGA	ACATGCT	CACGG	AGTCTC	ACAGTGC	GTGCTGGT	TGGAATG	CGTTCACT	1860
Qy	1861	TTTGGC	AGCTTTG	CAACATG	CTGCTTGG	GGGAAATGG	GGCTTGGT	TGCCATG	CAAT	1920
Db	1861	TTTGGC	AGCTTTG	CAACATG	CTGCTTGG	GGGAAATGG	GGCTTGGT	TGCCATG	CAAT	1920
Qy	1921	GGTTTTT	TGTTCC	ATTTTGTCTT	CAGAA	TCATCTGCTTT	CCCTGCTTACT	CTGGCAG	CCCTG	1980
Db	1921	GGTTTTT	TGTTCC	ATTTTGTCTT	CAGAA	TCATCTGCTTT	CCCTGCTTACT	CTGGCAG	CCCTG	1980
Qy	1981	GAGCGTGG	GTCTCTGTG	AAATATTTCTG	AAAAATTTG	AAACGAAAG	CTCAATTTCT	TGAC		2040
Db	1981	GAGCGTGG	GTCTCTGTG	AAATATTTCTG	AAAAATTTG	AAACGAAAG	CTCAATTTCT	TGAC		2040
Qy	2041	CTGAAAG	TATCATTTTG	CTCTG	CCCTGCTGG	CCCTGAGCA	TGGCCGAG	TTCCCTG		2100
Db	2041	CTGAAAG	TATCATTTTG	CTCTG	CCCTGCTGG	CCCTGAGCA	TGGCCGAG	TTCCCTG		2100
Qy	2101	CTGGTGG	CAGCAAG	TATGCG	CGCTCCCTCT	CTGCTGCTT	TGCTTTGG	GGAGCCCC		2160
Db	2101	CTGGTGG	CAGCAAG	TATGCG	CGCTCCCTCT	CTGCTGCTT	TGCTTTGG	GGAGCCCC		2160
Qy	2161	AGCA	CCATGGG	TACATG	TGTCCTCAT	TTGCTCA	ATTTCCCTTTG	CTCATG	ATG	2220
Db	2161	AGCA	CCATGGG	TACATG	TGTCCTCAT	TTGCTCA	ATTTCCCTTTG	CTCATG	ATG	2220
Qy	2221	ACCA	TTGCC	TACACCA	AGCTCTACTG	CAATTTG	GACAA	GGGAGAC	CTTATTTGG	2280
Db	2221	ACCA	TTGCC	TACACCA	AGCTCTACTG	CAATTTG	GACAA	GGGAGAC	CTTATTTGG	2280
Qy	2281	GACTG	CTCTATG	TGTAA	AAACATTTG	CCCTGCTCTT	CACAA	CTGCA	ATCTTAACTGC	2340
Db	2281	GACTG	CTCTATG	TGTAA	AAACATTTG	CCCTGCTCTT	CACAA	CTGCA	ATCTTAACTGC	2340
Qy	2341	CCTGTG	GGCTTTCTT	GTCTTCTCT	TTTAA	AAACCTTTA	CA	TTTTAT	TCAGTCTCTG	2400
Db	2341	CCTGTG	GGCTTTCTT	GTCTTCTCT	TTTAA	AAACCTTTA	CA	TTTTAT	TCAGTCTCTG	2400
Qy	2401	ATTAAG	TTTTTAT	TCCTTCTG	TGGT	TAGTCC	CACTTCTG	CATGCTCA	ATCTCCCTTCTCTAC	2460
Db	2401	ATTAAG	TTTTTAT	TCCTTCTG	TGGT	TAGTCC	CACTTCTG	CATGCTCA	ATCTCCCTTCTCTAC	2460

QY 61 GGGGCGAGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCCAACACACTGTCTATGC 120
DB 109 GGGGCGAGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCCAACACACTGTCTATGC 169
QY 121 GAGCCCGAGCGGAGGATTTGCTCAGGGGTGAGCTGCTCCGACCTCTGGGCTCTCGGAGCTG 180
DB 169 GAGCCCGAGCGGAGGATTTGCTCAGGGGTGAGCTGCTCCGACCTCTGGGCTCTCGGAGCTG 228
QY 181 CTTCTCAACCTCAGGCTCTTCACTCTCTCACTAGACCTCAGTATGAACAACATCAGTCAG 240
DB 229 CTTCTCAACCTCAGGCTCTTCACTCTCTCACTAGACCTCAGTATGAACAACATCAGTCAG 288
QY 241 CTGCTCCCGAATCCCTCCGAGCTCCGCTTCTGAGGAGTTACGCTTTGGCGGAAAC 300
DB 289 CTGCTCCCGAATCCCTCCGAGCTCCGCTTCTGAGGAGTTACGCTTTGGCGGAAAC 348
QY 301 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCTTAAAGTTCTTATG 360
DB 349 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCTTAAAGTTCTTATG 408
QY 361 CTGCAGAAATATCAGCTAAGACACGTAACCCACAGAGCTCTGCGAATTTGCGAAGCCTT 420
DB 409 CTGCAGAAATATCAGCTAAGACACGTAACCCACAGAGCTCTGCGAATTTGCGAAGCCTT 468
QY 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAGTGGC 480
DB 469 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAGTGGC 528
QY 481 CTGCAATCCCTGAGGACCTCTGGCTGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 540
DB 529 CTGCAATCCCTGAGGACCTCTGGCTGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 588
QY 541 GCTTTTGAAGTTTATCGGCAATGCAAGCAATGACCTTTGGCCCTGGAACAAATACACCAC 600
DB 589 GCTTTTGAAGTTTATCGGCAATGCAAGCAATGACCTTTGGCCCTGGAACAAATACACCAC 648
QY 601 ATACCAGACTATGCTTTGGAAACCTCTCAGCTTGTGTAGTCTTACATCTCCATAACAAT 660
DB 649 ATACCAGACTATGCTTTGGAAACCTCTCAGCTTGTGTAGTCTTACATCTCCATAACAAT 708
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 709 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 768
QY 721 TTAATTTACAAATTAACCTTGATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 780
DB 769 TTAATTTACAAATTAACCTTGATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 828
QY 781 GAACTAGGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 829 GAACTAGGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 888
QY 841 CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
DB 889 CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 948
QY 901 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAAATGCTGCTCACAATACTGAA 960
DB 949 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAAATGCTGCTCACAATACTGAA 1008
QY 961 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGAGATC 1020
DB 1009 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGAGATC 1068
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAGTCTAGATCTGTCT 1080
DB 1069 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAGTCTAGATCTGTCT 1128
QY 1081 TACAACCTATTAGAGATTTTACCCAGTCTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
DB 1129 TACAACCTATTAGAGATTTTACCCAGTCTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1188
QY 1141 CTAAGACATATTAAGAAATCTACGAAATTAAGTTTGAACCTTTCCAGCAGTTGCTTAGCCTC 1200

DB 1189 CTAAGACATATTAAGAAATCTACGAAATTAAGTTTGAACCTTTCCAGCAGTTGCTTAGCCTC 1248
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
DB 1249 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1308
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCGTCTGCTCTTTTCTATAAAT 1320
DB 1309 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCGTCTGCTCTTTTCTATAAAT 1368
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
DB 1369 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1428
QY 1381 TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGAGTGTCTG 1440
DB 1429 TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGAGTGTCTG 1488
QY 1441 GCATTTGGAGTGTGTGAGAAATGCTTAACTCAATCAATGGAATAAAGGTGACAAAC 1500
DB 1489 GCATTTGGAGTGTGTGAGAAATGCTTAACTCAATCAATGGAATAAAGGTGACAAAC 1548
QY 1501 AGCAGTATGAGACGACCTTTCAATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
DB 1549 AGCAGTATGAGACGACCTTTCAATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1608
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1609 GACCTTGAAGATTTCTGCTTGAATTTGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1668
QY 1621 TGTTTCACTTTCCCGAGGCCCTTCAAAACCTGTGGAACACCTGTGATGCTGCTGATC 1680
DB 1669 TGTTTCACTTTCCCGAGGCCCTTCAAAACCTGTGGAACACCTGTGATGCTGCTGATC 1728
QY 1681 AGAATTTGAGTGTGGACATAGCAGATTTCTGGCACTTACTTTGTAATGCTTTTGGTGAATCA 1740
DB 1729 AGAATTTGAGTGTGGACATAGCAGATTTCTGGCACTTACTTTGTAATGCTTTTGGTGAATCA 1788
QY 1741 ACAGTTTTCAGATCCCTCTGTATATTTCCCACTTAAATGTTAAATGGGGTTCATCGCA 1800
DB 1789 ACAGTTTTCAGATCCCTCTGTATATTTCCCACTTAAATGTTAAATGGGGTTCATCGCA 1848
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1849 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1908
QY 1861 TTTGGCAGCTTTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1909 TTTGGCAGCTTTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1969 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 2028
QY 1981 GAGCGTGGTTCCTGCTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2040
DB 2029 GAGCGTGGTTCCTGCTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2088
QY 2041 CTGAAAGTAAATCAATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2089 CTGAAAGTAAATCAATTTTGTCTCTGTGCCCTTGACCTTGACCAATGCGCGAGTTTCCCTG 2148
QY 2101 CTGGGTGGCAGCAGTATGGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2149 CTGGGTGGCAGCAGTATGGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208
QY 2161 AGCACCATGGCTACATGCTGCTCTCATCTTGCTCAATTTCCCTTCTGCTTCTCATGATG 2220
DB 2209 AGCACCATGGCTACATGCTGCTCTCATCTTGCTCAATTTCCCTTCTGCTTCTCATGATG 2268
QY 2221 ACCATTCCTTACCAACAGGCTCTACTGCAATTTGGAAGGGAGACCTTGGAGAAATATTGG 2280

|||||
Db 441 CTGCTCCCGAATCCCTCTGCCAGTCTCGCTTCTCGAGAGTTACGTCTTGGGGAAAC 500
Qy 301 GCTCTGACATACATTTCCAAAGGAGCATTCACGTGGCCCTTTACAGTCTTAAAGTTCTTATG 360
Db 501 GCTCTGACATACATTTCCAAAGGAGCATTCACGTGGCCCTTTACAGTCTTAAAGTTCTTATG 560
Qy 361 CTGCAGAAATACAGCTTAAGACACGTACCCACAGAGCTCTGAGAAATTTGCGAAGCCTT 420
Db 561 CTGCAGAAATACAGCTTAAGACACGTATACCCACAGAGCTCTGAGAAATTTGCGAAGCCTT 620
Qy 421 CAATCCCTGGCTCTGGATGCTTAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 621 CAATCCCTGGCTCTGGATGCTTAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 680
Qy 481 CTGCATTTCCCTGAGGCACTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCAG 540
Db 681 CTGCATTTCCCTGAGGCACTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCAG 740
Qy 541 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGACAAATACACCAC 600
Db 741 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGACAAATACACCAC 800
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 801 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 860
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCCTAGAGACTTTAGAT 720
Db 861 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCCTAGAGACTTTAGAT 920
Qy 721 TTTAAATTTACAAATACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCACCTTAAA 780
Db 921 TTTAAATTTACAAATACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCACCTTAAA 980
Qy 781 GAATAGGATTTTATAGCAACAATATCAGGTTCGATACCTCAGAAAGCAATTTGTAGGCAAC 840
Db 981 GAATAGGATTTTATAGCAACAATATCAGGTTCGATACCTCAGAAAGCAATTTGTAGGCAAC 1040
Qy 841 CCTTCTCTTAATTAACAATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT 900
Db 1041 CCTTCTCTTAATTAACAATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT 1100
Qy 901 TTTCAACATTTTACCTGAATTAAGAACACTGACTCTGAAATGGTGCCTCACAATAACTGAA 960
Db 1101 TTTCAACATTTTACCTGAATTAAGAACACTGACTCTGAAATGGTGCCTCACAATAACTGAA 1160
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 1161 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
Qy 1021 TCATCTCTCTTAACCCGCTCTGCAATCAGTTACCTTAATCTCCAGTGTAGATCTGTCT 1080
Db 1221 TCATCTCTCTCTTAACCCGCTCTGCAATCAGTTACCTTAATCTCCAGTGTAGATCTGTCT 1280
Qy 1081 TACAACTTATTAAGATTTACCCAGTTTTTACGTCTGCGCAAGCTTTCAGAAATTTGAC 1140
Db 1281 TACAACTTATTAAGATTTACCCAGTTTTTACGTCTGCGCAAGCTTTCAGAAATTTGAC 1340
Qy 1141 CTAAGACATTAATGAAATTTACGAAATTTAAAGTTGACACTTTCCAGCAAGTTGCTTAGCCTC 1200
Db 1341 CTAAGACATTAATGAAATTTACGAAATTTAAAGTTGACACTTTCCAGCAAGTTGCTTAGCCTC 1400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTAATCAACCAATGCAATTTTCACT 1260
Db 1401 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTAATCAACCAATGCAATTTTCACT 1460
Qy 1261 TTGCACTCCCTTAATAAGCTGGAACCTATCGTCCAACTCTGCTGCTTTTCTTATTAACCT 1320
Db 1461 TTGCACTCCCTTAATAAGCTGGAACCTATCGTCCAACTCTGCTGCTTTTCTTATTAACCT 1520
Qy 1321 GGGTTACATGGTTTAACTCACTTTAAAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
|||||

Db 1521 GGGTTACATGGTTTAACTCACTCACTTTAAAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1580
Qy 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTTACCAAGTGTGT 1440
Db 1581 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTTACCAAGTGTGT 1640
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCTATAGATTTCTAAATCAATGGAATAAAGGTGACAAAC 1500
Db 1641 GCATTTGGAGTGTGTGAGAAATGCTATAGATTTCTAAATCAATGGAATAAAGGTGACAAAC 1700
Qy 1501 AGCAGTATGGAACGACCTTTCAATAAGAAAGATGCTGGAAATTTTCAAGGCTCAAGATGAACGT 1560
Db 1701 AGCAGTATGGAACGACCTTTCAATAAGAAAGATGCTGGAAATTTTCAAGGCTCAAGATGAACGT 1760
Qy 1561 GACCTTTGAAGATTTTCCCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 1761 GACCTTTGAAGATTTTCCCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1820
Qy 1621 TGTTCACCTTTCCCGAGGCCCTTTCAAAACCTGTGAAACACCTGCTTGATGGCTGGCTGATC 1680
Db 1821 TGTTCACCTTTCCCGAGGCCCTTTCAAAACCTGTGAAACACCTGCTTGATGGCTGGCTGATC 1880
Qy 1681 AGAATTTGGAGTGTGGAACCATPAGCAGTTCTGGCACTTACTTTGTAAATGCTTTGGTGAATCA 1740
Db 1881 AGAATTTGGAGTGTGGAACCATPAGCAGTTCTGGCACTTACTTTGTAAATGCTTTGGTGAATCA 1940
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAAATTTGGGGTCAATCGCA 1800
Db 1941 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAAATTTGGGGTCAATCGCA 2000
Qy 1801 GCAGTGAACATGCTCAACGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGGAATGGCTTCACT 1860
Db 2001 GCAGTGAACATGCTCAACGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGGAATGGCTTCACT 2060
Qy 1861 TTTGGCAGCTTTTGCACGACATGCTGCTGGTGGGAGAAATGGGTTGGTTCGCCATGTCAAT 2120
Db 2061 TTTGGCAGCTTTTGCACGACATGCTGCTGGTGGGAGAAATGGGTTGGTTCGCCATGTCAAT 2180
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 2121 GGTTTTTTGTCCATTTTGTCTTCCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2180
Qy 1981 GACGTGGGTCTCTCTGTGAAATATTCTGCAAAATTTTGAACAGAAAGCTCAATTTCTTAGC 2040
Db 2181 GACGTGGGTCTCTCTGTGAAATATTCTGCAAAATTTTGAACAGAAAGCTCAATTTCTTAGC 2240
Qy 2041 CTGAAAGTAAATCAATTTTGTCTCTGCTGCTGCTGGCTGTCACATGCGCGCAGTTTCCCTG 2100
Db 2241 CTGAAAGTAAATCAATTTTGTCTCTGCTGCTGCTGGCTGTCACATGCGCGCAGTTTCCCTG 2300
Qy 2101 CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGGGAGCC 2160
Db 2301 CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGGGAGCC 2360
Qy 2161 AGCACCATGGGCTACATGCTGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCTCATGATG 2220
Db 2361 AGCACCATGGGCTACATGCTGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCTCATGATG 2420
Qy 2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTTGGAGAGACCTGGAGAAATATTTGG 2280
Db 2421 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTTGGAGAGACCTGGAGAAATATTTGG 2480
Qy 2281 GACTGCTCTATGTTAAACACATTTGCCCTGCTCTTCAACCAAGCTGCAATTTTAAACCTGC 2340
Db 2481 GACTGCTCTATGTTAAACACATTTGCCCTGCTCTTCAACCAAGCTGCAATTTTAAACCTGC 2540
Qy 2341 CCTGTGGCTTTCTGTCTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
Db 2541 CCTGTGGCTTTCTGTCTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA 2600
Qy 2401 ATTAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTGCTGATGTCTCAATCCCTTCTCTAC 2460
Db 2601 ATTAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTGCTGATGTCTCAATCCCTTCTCTAC 2660
|||||

QY 1081 TACAACCTATTAGAAGATTACCCAGTTTTTTCAGTCTGCCAAGGTTTCAGAAATTGAC 1140
DB 1281 TACAACCTATTAGAAGATTACCCAGTTTTTTCAGTCTGCCAAGGTTTCAGAAATTGAC 1340
QY 1141 CTAAGACATAATGAATCTACCAAAATTAAGTTGACACTTTCCAGCAGTTGCTAGCCTC 1200
DB 1341 CTAAGACATAATGAATCTACCAAAATTAAGTTGACACTTTCCAGCAGTTGCTAGCCTC 1400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTCGTAATTTACCCCAATGCAATTTCCACT 1260
DB 1401 CGATCGCTGAATTTGGCTTGGAAACAAAATTCGTAATTTACCCCAATGCAATTTCCACT 1460
QY 1261 TTGCCATCGCTAATAAGCTGACACTATGCTCAACCTCTGCTGCTTTTCCCTAATCT 1320
DB 1461 TTGCCATCGCTAATAAGCTGACACTATGCTCAACCTCTGCTGCTTTTCCCTAATCT 1520
QY 1321 GGGTTACATGGTTTAACTCACITTAATAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
DB 1521 GGGTTACATGGTTTAACTCACITTAATAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1580
QY 1381 TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGCTTTATGCTTTACAGTGTGT 1440
DB 1581 TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGCTTTATGCTTTACAGTGTGT 1640
QY 1441 GCATTTGGAGTGTGAGNAATGCTTAAAGATTTCTAATCAATGGAATAAGGTCACAC 1500
DB 1641 GCATTTGGAGTGTGAGNAATGCTTAAAGATTTCTAATCAATGGAATAAGGTCACAC 1700
QY 1501 ASCAGTATGAGACGCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAAGT 1560
DB 1701 ASCAGTATGAGACGCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAAGT 1760
QY 1561 GACCTTGAAGATTTCTGCTTGCATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB 1761 GACCTTGAAGATTTCTGCTTGCATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAG 1820
QY 1621 TGTTTCACTTCCAGGCCCCCTTCAACCTGTGACACCTCTGTGAGGCTGGCTGATC 1680
DB 1821 TGTTTCACTTCCAGGCCCCCTTCAACCTGTGACACCTCTGTGAGGCTGGCTGATC 1880
QY 1681 AGAATTTGAGTGTGGACCATAGCAGTCTGGCAGCTTACCTTGTAAATGCTTTGTGACTTCA 1740
DB 1881 AGAATTTGAGTGTGGACCATAGCAGTCTGGCAGCTTACCTTGTAAATGCTTTGTGACTTCA 1940
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGCTCATGCA 1800
DB 1941 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGCTCATGCA 2000
QY 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGGCTTCCT 1860
DB 2001 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGGCTTCCT 2060
QY 1861 TTTGGCAGCTTTGCACGACATGGTGTGCTGGTGGGAGATGGGTTGGTTGCCATGCTCAAT 1920
DB 2061 TTTGGCAGCTTTGCACGACATGGTGTGCTGGTGGGAGATGGGTTGGTTGCCATGCTCAAT 2120
QY 1921 GGTTTTTTGTCCATTTTGCCTTCAGAAATCATCTGTTTTCCTGCTTACTCTGGAGCCCTG 1980
DB 2121 GGTTTTTTGTCCATTTTGCCTTCAGAAATCATCTGTTTTCCTGCTTACTCTGGAGCCCTG 2180
QY 1981 GAGCGTGGGTTCTGTGAAATTAATCTGCAAAATTTGAACAGAAAGCTCCATTTTCTAGC 2040
DB 2181 GAGCGTGGGTTCTGTGAAATTAATCTGCAAAATTTGAACAGAAAGCTCCATTTTCTAGC 2240
QY 2041 CTGAAGATTAATCAATTTGCTGTGCTGCTGCTGGCTTGCACCATGGCCGAGTTCCCTG 2100
DB 2241 CTGAAGATTAATCAATTTGCTGTGCTGCTGCTGGCTTGCACCATGGCCGAGTTCCCTG 2300
QY 2101 CTGGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2160
DB 2301 CTGGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2360

QY 2161 AGCACCATGGGCTACATGGTGTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
DB 2361 AGCACCATGGGCTACATGGTGTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2420
QY 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAAGGGAGAGCTGGAGAATATTGG 2280
DB 2421 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAAGGGAGAGCTGGAGAATATTGG 2480
QY 2281 GACTGCTCTATGGTAAAAACACATTGGCCCTGTGCTCTTACCAACCTGATCTTAACTGC 2340
DB 2481 GACTGCTCTATGGTAAAAACACATTGGCCCTGTGCTCTTACCAACCTGATCTTAACTGC 2540
QY 2341 CCTGGGCTTCTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAAGTCTCTGAAGTA 2400
DB 2541 CCTGGGCTTCTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAAGTCTCTGAAGTA 2600
QY 2401 ATTAAGTTTATCTCTTCTGGTGTAGTCCACCTCTCTGCTGATGTCTCAATCCCTTCTCTAC 2460
DB 2601 ATTAAGTTTATCTCTTCTGGTGTAGTCCACCTCTCTGCTGATGTCTCAATCCCTTCTCTAC 2660
QY 2461 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2520
DB 2661 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2720
QY 2521 TGGACAAGATCAAAACACCCCAAGCTTGTATGCAATTAATCTCTGATGATGTGAAAAACAG 2580
DB 2721 TGGACAAGATCAAAACACCCCAAGCTTGTATGCAATTAATCTCTGATGATGTGAAAAACAG 2780
QY 2581 TCCTGTGATCTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
DB 2781 TCCTGTGATCTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTG 2840
QY 2641 CTTCCAGTCTGGTGGCATCACAGCTTATCCAGTGTAGAGCTGCCATCTTCTCTCT 2700
DB 2841 CTTCCAGTCTGGTGGCATCACAGCTTATCCAGTGTAGAGCTGCCATCTTCTCTCTCT 2900
QY 2701 GTGGCATTTGTCCCATGTCTCTAA 2724
DB 2901 GTGGCATTTGTCCCATGTCTCTAA 2924

RESULT 11

US-10-505-486-205
; Sequence 205, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 205
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Human
US-10-505-486-205

Query Match 98.0%; Score 2670; DB 21; Length 3438;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2720; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGACACTCCCGCTCGGTGTGCTCTGCTTCCCTTGTCTGTCTGCTGAGCTGGCGACC 60
DB 1 ATGGACACTCCCGCTCGGTGTGCTCTGCTTCCCTTGTCTGTCTGCTGAGCTGGCGACC 60
QY 61 GGGGGCAGCTCTCTCCAGGTCTGGTGTGTTGTGAGGGGCTGCCCAACACACTGTCTATTC 120

Db 61 GGGGGGAGCTCTCCAGGCTCGTGTTGCTGAGGGGCTGCCCAACACACTGTCAATGC 120
Qy 121 GAGCCCGAGCGCAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG 180
Db 121 GAGCCCGAGCGCAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG 180
Qy 181 CTTTCCAACTCAGCGCTTTCACCTCTACCTTAGACTCTGATGAAACAACTCAGTCAG 240
Db 181 CTTTCCAACTCAGCGCTTTCACCTCTACCTTAGACTCAGTATGAAACAACTCAGTCAG 240
Qy 241 CTGCTCCGAACTCCCTGCCAGTCTCCGCTTCTCGAGGAGTACGTCTTGGGGAAAC 300
Db 241 CTGCTCCGAACTCCCTGCCAGTCTCCGCTTCTCGAGGAGTACGTCTTGGGGAAAC 300
Qy 301 GCTCTGACATACATTTCCAGGGAGCAATTCACCTGGCTTTACAGTCTTAAAGTTCATTG 360
Db 301 GCTCTGACATACGTTCCCAAGGGAGCAATTCACCTGGCTTTACAGTCTTAAAGTTCATTG 360
Qy 361 CTGCAGAAATACAGCTTAGACACGTAACCAAGAGCTCTGAGAAATTTGGAGGCTT 420
Db 361 CTGCAGAAATACAGCTTAGACACGTAACCAAGAGCTCTGAGAAATTTGGAGGCTT 420
Qy 421 CAATCCCTGCTGATGCTAACCAACATCAGCTATGTGCCCCCAAGCTGTTTCACTGGC 480
Db 421 CAATCCCTGCTGATGCTAACCAACATCAGCTATGTGCCCCCAAGCTGTTTCACTGGC 480
Qy 481 CTGCATTTCCCTGAGGACCTGCTGGCTGATGCAATGCGTTTACAGAAATCCCGCTCCAG 540
Db 481 CTGCATTTCCCTGAGGACCTGCTGGCTGATGCAATGCGTTTACAGAAATCCCGCTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCAATGCAAGCATGACCTTGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTGAAGTTTATCGGCAATGCAAGCATGACCTTGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTGGAACTCTCCAGCTTGGTAGTTTACATCTCCATACAAAT 560
Db 601 ATACAGACTATGCTTTGGAACTCTCCAGCTTGGTAGTTTACATCTCCATACAAAT 560
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTATGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTATGAT 720
Qy 721 TTAATAACAATAA CTTTGAATTTCCCACTGCAATTAGAGACTCTCCAACTTAAA 780
Db 721 TTAATAACAATAA CTTTGAATTTCCCACTGCAATTAGAGACTCTCCAACTTAAA 780
Qy 781 GAATAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGCAAC 840
Db 781 GAATAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGCAAC 840
Qy 841 CTTTCTCTTATTAACAATAATTTCTATGCAATCCCATCCAATTTGTGGAGATCTGCT 900
Db 841 CTTTCTCTTATTAACAATAATTTCTATGCAATCCCATCCAATTTGTGGAGATCTGCT 900
Qy 901 TTTCAACAATTTACCTGAACTAGAAACACTGACTCTGAAATGGTGCCTCACAATAAATGAA 960
Db 901 TTTCAACAATTTACCTGAACTAGAAACACTGACTCTGAAATGGTGCCTCACAATAAATGAA 960
Qy 961 TTTCTGATTTAACTGGAACTGAAACCTGGAGAGTCTGACTTAACTGAGGACAGATC 1020
Db 961 TTTCTGATTTAACTGGAACTGAAACCTGGAGAGTCTGACTTAACTGAGGACAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAGTGTAGATCTGCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAGTGTAGATCTGCT 1080
Qy 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCACTGCTGCCAAAGCTTACAGAAATTTGAC 1140
Db 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCACTGCTGCCAAAGCTTACAGAAATTTGAC 1140
Qy 1141 CTAAAGACATTAAGAAATCTACGAAATTTAAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1200

Db 1141 CTAAAGACATTAAGAAATCTACGAAATTTAAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAAGGAAATCAATGCTTTACAGACTTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAAGGAAATCAATGCTTTACAGACTTTGATA 1380
Qy 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATCCAGTCTCTGT 1440
Db 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATCCAGTCTCTGT 1440
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATTAAGGTGACAC 1500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATTAAGGTGACAC 1500
Qy 1501 AGCAGTATGAGACACTTTCATAGAAAGATGCTGGAATGTTTCCAGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGAGACACTTTCATAGAAAGATGCTGGAATGTTTCCAGCTCAAGATGAACGT 1560
Qy 1561 GACCTTTGAAGATTTTCTGCTTTGAGGAGACCTGAAAGCCCTTCAATCCAGTGCAG 1620
Db 1561 GACCTTTGAAGATTTTCTGCTTTGAGGAGACCTGAAAGCCCTTCAATCCAGTGCAG 1620
Qy 1621 TGTTCACTTTCCCGAGGCCCTTCAAACTGTTGAACACTGCTGTTGATGGCTGGCTGATC 1680
Db 1621 TGTTCACTTTCCCGAGGCCCTTCAAACTGTTGAACACTGCTGTTGATGGCTGGCTGATC 1680
Qy 1681 AGAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
Db 1681 AGAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
Qy 1741 ACAGTTTTTCCAGATCCCTCTGTACATTTCCCACTTTAAACTGTTAAATGGGGTCACTGCA 1800
Db 1741 ACAGTTTTTCCAGATCCCTCTGTACATTTCCCACTTTAAACTGTTAAATGGGGTCACTGCA 1800
Qy 1801 GCAGTGAACTGCTCAGGGAGTCTCAGTGGCTGCTGGCTGGTGGATGGTGGTCACT 1860
Db 1801 GCAGTGAACTGCTCAGGGAGTCTCAGTGGCTGCTGGCTGGTGGATGGTGGTCACT 1860
Qy 1861 TTTTGGCAGCTTTGCAAGCATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 1920
Db 1861 TTTTGGCAGCTTTGCAAGCATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTTCCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGGTTCTGTGAAATATTTGCAAAATTTGAAAAGCTCAATTTTCTAGC 2040
Db 1981 GAGCGTGGGTTCTGTGAAATATTTGCAAAATTTGAAAAGCTCAATTTTCTAGC 2040
Qy 2041 CTGAAAGTAACTAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CTGAAAGTAACTAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTTGTGCTTTGGGGAGCCC 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTTGTGCTTTGGGGAGCCC 2160
Qy 2161 AGCAACATGGGTTAATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGTCTCTCATGATG 2220
Db 2161 AGCAACATGGGTTAATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGTCTCTCATGATG 2220
Qy 2221 ACCATTTGCCCTACCAAGCTCTACGCAATTTGGAAAGGGAGACTGGAGAAATTTTGG 2280
Db 2221 ACCATTTGCCCTACCAAGCTCTACGCAATTTGGAAAGGGAGACTGGAGAAATTTTGG 2280

QY	2281	GA	CTGCTCTATGGTAAACACATTGGCCCTGTGCTCTTACCAACCTGCAATCTAAACTGC	2340
DB	2281	GA	CTGCTCTATGGTAAACACATTGGCCCTGTGCTCTTACCAACCTGCAATCTAAACTGC	2340
QY	2341	CT	GTGGGCTTCTGTGCTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA	2400
DB	2341	CT	GTGGGCTTCTGTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA	2400
QY	2401	AT	TAAAGTTTATCTCTGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCCCTTCTCTAC	2460
DB	2401	AT	TAAAGTTTATCTCTGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCCCTTCTCTAC	2460
QY	2461	AT	CTTGTTCATCTCTTAAAGGAGATCTGGTAGCCTGAGAAAGCAACCTTACGTC	2520
DB	2461	AT	CTTGTTCATCTCTTAAAGGAGATCTGGTAGCCTGAGAAAGCAACCTTACGTC	2520
QY	2521	TG	GACAAGATCAAAACACCAAGCTTGATGTCATTTAACTCTGTGATGATGTCGAAAAACAG	2580
DB	2521	TG	GACAAGATCAAAACACCAAGCTTGATGTCATTTAACTCTGTGATGATGTCGAAAAACAG	2580
QY	2581	TC	CTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCATTTATGACCTG	2640
DB	2581	TC	CTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCATTTATGACCTG	2640
QY	2641	CT	CCAGTTCGGTGCATCAGAGCTTATCCAGTGATGAGAGCTGCCATCTTTCTCT	2700
DB	2641	CT	CCAGTTCGGTGCATCAGAGCTTATCCAGTGATGAGAGCTGCCATCTTTCTCT	2700
QY	2701	GT	GGCAATTTGTCCTCATGCTC 2721	
DB	2701	GT	GGCAATTTGTCCTCATGCTC 2721	
RESULT 12				
US-10-251-385-277				
; Sequence 277, Application US/10251385				
; Publication No. US20030105292A1				
; GENERAL INFORMATION:				
; APPLICANT: Behan, Dominic P.				
; APPLICANT: Liaw, Chen W.				
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G				
; TITLE OF INVENTION: Protein-Coupled				
; FILE REFERENCE: AREN-0040				
; CURRENT APPLICATION NUMBER: US/10/251,385				
; CURRENT FILING DATE: 2002-09-20				
; PRIOR APPLICATION NUMBER: US/09/170,496				
; PRIOR FILING DATE: 1998-10-13				
; NUMBER OF SEQ ID NOS: 294				
; SOFTWARE: PatentIn version 3.1				
; SEQ ID NO 277				
; LENGTH: 2724				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-251-385-277				
Query Match 96.3%; Score 2622; DB 15; Length 2724;				
Best Local Similarity 99.9%; Pred. No. 0;				
Matches 2722; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	1	AT	GACACCTCCGGCTCGGTGTGCTCTCTGCTCTTCCCTGTGCTGTGCTGAGTGGGACC	60
DB	1	AT	GACACCTCCGGCTCGGTGTGCTCTCTGCTCTTCCCTGTGCTGTGCTGAGTGGGACC	60
QY	61	GG	GGGAGCTCTCCAGCTGTGGTGTGCTGAGGGGCTGCCACACACTGTCTTGC	120
DB	61	GG	GGGAGCTCTCTCCAGCTGTGGTGTGCTGAGGGGCTGCCACACACTGTCTTGC	120
QY	121	GAG	CCGAGGAGATGTGTGCTCAGGGTGGAGTGTCTCCGAGCTGGGGCTCTCGGAGCTG	180
DB	121	GAG	CCGAGGAGATGTGTGCTCAGGGTGGAGTGTCTCCGAGCTGGGGCTCTCGGAGCTG	180

QY	181	CT	TTCAACCTCAGCGTCTTTCACCTCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG	240
DB	181	CT	TTCAACCTCAGCGTCTTTCACCTCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG	240
QY	241	CT	GTCTCCCGAATCCCGCTCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTTCGGGGAAC	300
DB	241	CT	GTCTCCCGAATCCCGCTCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTTCGGGGAAC	300
QY	301	GC	CTGACATCATTTCCCAAGGAGCATTCATCTGGCCCTTTACAGTCTTAAAGTTCTTATG	360
DB	301	GC	CTGACATCATTTCCCAAGGAGCATTCATCTGGCCCTTTACAGTCTTAAAGTTCTTATG	360
QY	361	CT	GAGAAATATCAGTAAAGCACCTACCCACAGAGCTCTGCAGAAATTTGGGAAGCCTT	420
DB	361	CT	GAGAAATATCAGTAAAGCACCTACCCACAGAGCTCTGCAGAAATTTGGGAAGCCTT	420
QY	421	CA	ATCCCTGCGTCTCGATGTCTAAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
DB	421	CA	ATCCCTGCGTCTCGATGTCTAAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
QY	481	CT	GCAATCCCTGAGGCACTGTGTGCTGATGACAAATGGTTAAACAGAAATTCGCCGTCAG	540
DB	481	CT	GCAATCCCTGAGGCACTGTGTGCTGATGACAAATGGTTAAACAGAAATTCGCCGTCAG	540
QY	541	GC	TTTATAGAACTTATCGGCATTCGCAAGCCATGACCTTGGCCCTGAAACAAATATACCAAC	600
DB	541	GC	TTTATAGAACTTATCGGCATTCGCAAGCCATGACCTTGGCCCTGAAACAAATATACCAAC	600
QY	601	AT	ACAGACTATGCTTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT	660
DB	601	AT	ACAGACTATGCTTTTGGAAACCTCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT	660
QY	661	AGA	TCCACTCCCTGGGAAAGAAATGCTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT	720
DB	661	AGA	TCCACTCCCTGGGAAAGAAATGCTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT	720
QY	721	TT	TAATTAACATTAACCTTGTGATGAAATTCGCCCACTGCAATTAGGACACTCTCCAACTTAA	780
DB	721	TT	TAATTAACATTAACCTTGTGATGAAATTCGCCCACTGCAATTAGGACACTCTCCAACTTAA	780
QY	781	GA	ACTAGGATTTATAGCAAAATATCAGTGTGATGAGGCTCCACAGCCTAGAGACTTTAGGCAAC	840
DB	781	GA	ACTAGGATTTATAGCAAAATATCAGTGTGATGAGGCTCCACAGCCTAGAGACTTTAGGCAAC	840
QY	841	CT	CTCTTTATCAATCATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT	900
DB	841	CT	CTCTCTTTATCAATCATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT	900
QY	901	TT	TCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCACAATAAATCTGAA	960
DB	901	TT	TCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCACAATAAATCTGAA	960
QY	961	TT	TCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020	
DB	961	TT	TCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020	
QY	1021	TC	ATCTCTTCTCAAAACCGTCTGCAATCAGTTAATCTCCAAAGTGTAGATCTGTCT 1080	
DB	1021	TC	ATCTCTTCTCAAAACCGTCTGCAATCAGTTAATCTCCAAAGTGTAGATCTGTCT 1080	
QY	1081	TAC	ACCTATTTAGAGATTTACCCAGTTTTCAGTCTGCCAAAGCTCCAGAAATTTGAC 1140	
DB	1081	TAC	ACCTATTTAGAGATTTACCCAGTTTTCAGTCTGCCAAAGCTCCAGAAATTTGAC 1140	
QY	1141	CT	AAGACATTAATGAAATCTACGAAATTAAGTTGACATCTTCCAGCAGTGTCTTACGCTTC	1200
DB	1141	CT	AAGACATTAATGAAATCTACGAAATTAAGTTGACATCTTCCAGCAGTGTCTTACGCTTC	1200
QY	1201	CG	ATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260
DB	1201	CG	ATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260

QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTGTCGTCTTTCTCTAATACT 1320
DB 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTGTCGTCTTTCTCTAATACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATAAAGGAAATCATGCCCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGGTTTAACTCACTTAAATAAAGGAAATCATGCCCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
QY 1441 GCATTTGGAGTGTGAGATGCTATAGATTTCTAAATCAATGGAATAAGGTGACAAC 1500
DB 1441 GCATTTGGAGTGTGAGATGCTATAGATTTCTAAATCAATGGAATAAGGTGACAAC 1500
QY 1501 AGCAGTATGACGACCTTCAATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
DB 1501 AGCAGTATGACGACCTTCAATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
QY 1561 GACCTTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB 1561 GACCTTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY 1621 TGTTCACCTTCCCGAGCCCTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACCTTCCCGAGCCCTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
QY 1681 AGAATTTGGAGTGTGGAACCATAGCAGTTCTGGCACTTACTTGTATGCTTTGGTGACTTCA 1740
DB 1681 AGAATTTGGAGTGTGGAACCATAGCAGTTCTGGCACTTACTTGTATGCTTTGGTGACTTCA 1740
QY 1741 ACAGTTTTCAGATPCCCTCTGTACATTTCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
DB 1741 ACAGTTTTCAGATPCCCTCTGTACATTTCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
QY 1801 GCAGTGAACATGCTCAGGAGTCTCAGTGCCTGCTGCTGGTGTGATGATGCTTCACT 1860
DB 1801 GCAGTGAACATGCTCAGGAGTCTCAGTGCCTGCTGCTGGTGTGATGATGCTTCACT 1860
QY 1861 TTTGGCAGCTTTGCAGCAGATGCTGCTGCTGGGAGAAATGGGTTGGTCCCATGTCATT 1920
DB 1861 TTTGGCAGCTTTGCAGCAGATGCTGCTGCTGGGAGAAATGGGTTGGTCCCATGTCATT 1920
QY 1921 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1921 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
QY 1981 GAGCGTGGGTTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 1981 GAGCGTGGGTTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAAATCAATTTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGAAAGTAAATCAATTTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACATGGGTATCAATGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
DB 2161 AGCACATGGGTATCAATGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
QY 2221 ACCATTTGCTTACACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
DB 2221 ACCATTTGCTTACACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
QY 2281 GACTGCTCTATGTAACCAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
DB 2281 GACTGCTCTATGTAACCAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
QY 2341 CCTGNGGCTTTCTGCT 2400

RESULT 13

US-10-295-027-483
; Sequence 483, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 483
; LENGTH: 2651

Query Match	Best Local Similarity	Matches	Score	DB	Length	Indels	Gaps	0;
857	TACATTTCTATGACAAATCCCATCTTATTTGTTGGAGATCTCTTTTCAACATTTTACCTG	916	68.4%;	DB 17;	2651;	0;	0;	0;
785	TACATTTCTATGACAAATCCCATCTTATTTGTTGGAGATCTCTTTTCAACATTTTACCTG	844						
917	AACTAAGAACACATGACCTGAAATGCTGCTCAAAATAACATTAATCTCTGATTTAACTG	976						
845	AACTAAGAACACATGACCTGAAATGCTGCTCAAAATAACATTAATCTCTGATTTAACTG	904						
977	GAACCTGCAACCTGGAGAGTCTGACCTTAACTTGAAGCAGAGATCTCATCTCTCTCTCAAA	1036						
905	GAACCTGCAACCTGGAGAGTCTGACCTTAACTTGAAGCAGAGATCTCATCTCTCTCTCAAA	964						
1037	CCGCTGCTCAATCAGTTTACCTTAACTTCAAGTGTCTAGATCTGCTTACCAACCTATTAGAAG	1096						
965	CCGCTGCTCAATCAGTTTACCTTAACTTCAAGTGTCTAGATCTGCTTACCAACCTATTAGAAG	1024						
1097	ATTTCACCACTTTTTCAGTCTGCAAAAGCTTCAGAAATTAACCTAAGACATTAATGAAA	1156						
1025	ATTTCACCACTTTTTCAGTCTGCAAAAGCTTCAGAAATTAACCTAAGACATTAATGAAA	1084						
1157	TCTACGAAATTAAGTTGACATTTTCAGAGTTCCTTAGCTTCGATCGCTGAATTTGG	1216						
1085	TCTACGAAATTAAGTTGACATTTTCAGAGTTCCTTAGCTTCGATCGCTGAATTTGG	1144						
1217	CTTGGAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTTGCCATCCCTTAATA	1276						
1145	CTTGGAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTTGCCATCCCTTAATA	1204						
1277	AGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTATATACTGGGTTTACATGGTTTAA	1336						
1205	AGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTATATACTGGGTTTACATGGTTTAA	1264						
1337	CTCATTTAAATTTAAACAGGAATCATGCTTACAGAGCTTGATATCATCTGAAACTTTTC	1396						
1265	CTCATTTAAATTTAAACAGGAATCATGCTTACAGAGCTTGATATCATCTGAAACTTTTC	1324						
1397	CAGAACTCAAGGTTTATAGAAATGCTTATGCTTACAGAGCTTGATATCATCTGAAAGTGTG	1456						
1325	CAGAACTCAAGGTTTATAGAAATGCTTATGCTTACAGAGCTTGATATCATCTGAAAGTGTG	1384						
1457	AGAATGCTTATAAGATTTCTTAATCAATGGAATTAAGGTGACAAACAGCAGTATGGACGACC	1516						
1385	AGAATGCTTATAAGATTTCTTAATCAATGGAATTAAGGTGACAAACAGCAGTATGGACGACC	1444						
1517	TTCAATAAGAAAGATGCTGGAATGTTTACAGCTCAAGATGAAGTGAACCTTGAAGATTTCC	1576						
1445	TTCAATAAGAAAGATGCTGGAATGTTTACAGCTCAAGATGAAGTGAACCTTGAAGATTTCC	1504						
1577	TGCTTGACTTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCCAG	1636						
1505	TGCTTGACTTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCCAG	1564						
1637	GCCCTTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATCAGAAATTTGAGTGTGGA	1696						
1565	GCCCTTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATCAGAAATTTGAGTGTGGA	1624						
1697	CCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTTGGTGTGACTTCAACAGTTTTCAGATCCC	1756						
1625	CCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTTGGTGTGACTTCAACAGTTTTCAGATCCC	1684						
1757	CTCTGTACATTTCCCACTTAATGCTTTTGGGTCATCCGACGAGTGAACATGCTCA	1816						
1685	CTCTGTACATTTCCCACTTAATGCTTTTGGGTCATCCGACGAGTGAACATGCTCA	1744						
1817	CGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGATGCTGTTCACTTTTGGCAGCTTTGAC	1876						

; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1113
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-295-027-1113

Query Match 68.4%; Score 1864; DB 17; Length 2651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 857 TACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTTACCTG 916
Db 785 TACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTTACCTG 844

Qy 917 AACTAAGAACACTGACTCTGAAATGCTGCTCAAAATTAACGAATTTCTGATTTAACTG 976
Db 845 AACTAAGAACACTGACTCTGAAATGCTGCTCAAAATTAACGAATTTCTGATTTAACTG 904

Qy 977 GAACCTGCAAACTGGAGAGTCTGACTTTAACTGGAGACAGATCTCATCTCTCTCTCAA 1036
Db 905 GAACCTGCAAACTGGAGAGTCTGACTTTAACTGGAGACAGATCTCATCTCTCTCTCAA 964

Qy 1037 CCGTCTGCAATCAGTTACTTAATCTCAAGTGTAGATCTGTCTTACAACTATTAGAA 1096
Db 965 CCGTCTGCAATCAGTTACTTAATCTCAAGTGTAGATCTGTCTTACAACTATTAGAA 1024

Qy 1097 ATTTACCAAGTTTTTTCAGTCTGCGCAAAAGCTTTCAGAAATTTGACTAAGACATATGAAA 1156
Db 1025 ATTTACCAAGTTTTTTCAGTCTGCGCAAAAGCTTTCAGAAATTTGACTAAGACATATGAAA 1084

Qy 1157 TCTAGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGG 1216
Db 1085 TCTAGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGG 1144

Qy 1217 CTTGGAACAAATTTGCTATTATTTACCCCAATGCAATTTTCCACTTTGCGATCCCTTAATA 1276
Db 1145 CTTGGAACAAATTTGCTATTATTTACCCCAATGCAATTTTCCACTTTGCGATCCCTTAATA 1204

Qy 1277 AGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTTATAACTGGGTTACATGGTTTAA 1336

Db 1205 AGCTGGACCTATCGTCCAACTCTCTGCTGCTTTTCTTATAACTGGGTTACATGGTTTAA 1264
Qy 1337 CTCACCTTAAAAATTAAACAGGAATCATGCTCTTACAGAGCTTGATATCATCTGAAAACTTTTC 1396
Db 1265 CTCACCTTAAAAATTAAACAGGAATCATGCTCTTACAGAGCTTGATATCATCTGAAAACTTTTC 1324
Qy 1397 CAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTG 1456
Db 1325 CAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTG 1384
Qy 1457 AGAATGCTCTATAAGATTTCTAATCAATGGAATAAGAGTGACAAACAGCAGTATGGACGACC 1516
Db 1385 AGAATGCTCTATAAGATTTCTAATCAATGGAATAAGAGTGACAAACAGCAGTATGGACGACC 1444
Qy 1517 TTCTATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGCTGACCTTTGAAGATTTCC 1576
Db 1445 TTCTATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGCTGACCTTTGAAGATTTCC 1504
Qy 1577 TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCGAG 1636
Db 1505 TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCGAG 1564
Qy 1637 GCGCCTTCAAAACCTGTGAACACCTGCTTGTGATGGCTGCTGATCAGAAATTTGGAGTGTGA 1696
Db 1565 GCGCCTTCAAAACCTGTGTGAACACCTGCTTGTGATGGCTGCTGATCAGAAATTTGGAGTGTGA 1624
Qy 1697 CCATAGCAGTTTCTGGCAGTTACTTGTATGCTTGTGCTGCTTCAACAGTGTTCAGATCCC 1756
Db 1625 CCATAGCAGTTTCTGGCAGTTACTTGTATGCTTGTGCTGCTTCAACAGTGTTCAGATCCC 1684
Qy 1757 CTCTGTACATTTCCCAATTAACATGTTAAATTTGGGCTCATCGCAGCAGTGAACATGCTCA 1816
Db 1685 CTCTGTACATTTCCCAATTAACATGTTAAATTTGGGCTCATCGCAGCAGTGAACATGCTCA 1744
Qy 1817 CGGAGTCTCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876
Db 1745 CGGAGTCTCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1804
Qy 1877 GACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1936
Db 1805 GACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864
Qy 1937 TTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTG 1996
Db 1865 TTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTG 1924
Qy 1997 TGAATATTCTGCAAAATTTGAAACGAAGCTCCATTTTCTAGCTGAAAGTAAATCATTT 2056
Db 1925 TGAATATTCTGCAAAATTTGAAACGAAGCTCCATTTTCTAGCTGAAAGTAAATCATTT 1984
Qy 2057 TGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2116
Db 1985 TGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2044
Qy 2117 ATGGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGAGCCCAAGCAGTATGAGCTTACA 2176
Db 2045 ATGGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGAGCCCAAGCAGTATGAGCTTACA 2104
Qy 2177 TGCTGCTCTCATCTTGTCTGATTTCCCTTGTCTTCTCATGATGACCAATTTGGCTTACCA 2236
Db 2105 TGCTGCTCTCATCTTGTCTGATTTCCCTTGTCTTCTCATGATGACCAATTTGGCTTACCA 2164
Qy 2237 AGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATTTTGGGAGTGTCTCTATGGTAA 2296
Db 2165 AGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATTTTGGGAGTGTCTCTATGGTAA 2224
Qy 2297 AACACATTTGCTGCTTCTTCAACCACTGATCTTAACTGCTTAACTGCTGCTTCTTGT 2356
Db 2225 AACACATTTGCTGCTTCTTCAACCACTGATCTTAACTGCTTAACTGCTGCTTCTTGT 2284
Qy 2357 CTTTCTCTCTTTTAAATTAACCTTACATTTATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2416

Db	2385	CCTTCTCCTCTTTTAATAAACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCCTTC	2344
Qy	2417	TGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTCTTCTACATCTTGTTCAAATCCTC	2476
Db	2345	TGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTACATCTTGTTCAAATCCTC	2404
Qy	2477	ACTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAAAACCTACGTCTGGACAAGATCAAAAC	2536
Db	2405	ACTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAAAACCTACGTCTGGACAAGATCAAAAC	2464
Qy	2537	ACCAAGCTTGATGTCAATTAATCTCTGATGATGTGCAAAACAGTCTGTGACTCAACTC	2596
Db	2465	ACCAAGCTTGATGTCAATTAATCTCTGATGATGTGCAAAACAGTCTGTGACTCAACTC	2524
Qy	2597	AAGCCTTTGGTAAACCTTTACACAGCTCCACGATCACCTTATGACCTGCCCTCCAGTTTCCGTGC	2656
Db	2525	AAGCCTTTGGTAAACCTTTACACAGCTCCACGATCACCTTATGACCTGCCCTCCAGTTTCCGTGC	2584
Qy	2657	CATCACCAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTTCTGTGTGGCAATTTGTGCCCAT	2716
Db	2585	CATCACCAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTTCTGTGTGGCAATTTGTGCCCAT	2644
Qy	2717	GTCT 2720	
Db	2645	GTCT 2648	

RESULT 15

```

US-10-173-999-27
; Sequence 27, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-999-27

```

	Query Match	68.4%	Score 1864;	DB 17;	Length 2651;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1864;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	857	TACATTCTATGACAATCCCATCAATTTGTTGGGAGATCTGCTTTTCAACATTACTG	916		
Db	785	TACATTCTATGACAATCCCATCAATTTGTTGGGAGATCTGCTTTTCAACATTACTG	844		
Qy	917	AACTAAGAACTGACTCTGAATGGTGCTCAAAAATAACTGAATTTCTCATTTAACTG	976		
Db	845	AACTAAGAACTGACTCTGAATGGTGCTCAAAAATAACTGAATTTCTCATTTAACTG	904		
Qy	977	GAACTGCAAACTGGAGAGTCTGACTTTTAATCGAGCACAGATCTCATCTCTTCTCTCAAA	1036		
Db	905	GAACTGCAAACTGGAGAGTCTGACTTTTAATCGAGCACAGATCTCATCTCTCTCTCAAA	964		

```
|||||
Db 2045 ATGGCGCTCCCTCTCTGCGCTGCTTTTGGGAGCCAGCACCATGGGCTACA 2104
Qy 2177 TGGTCGCTCTCATCTTGTCTCAATCCCTTTGCTTCTCTCATGATGACCAATGGCTACACCA 2236
Db 2105 TGGTCGCTCTCACTTGTCTCAATCCCTTTGCTTCTCTCATGATGACCAATGGCTACACCA 2164
Qy 2237 AGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGATAA 2296
Db 2165 AGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGATAA 2224
Qy 2297 AACACATGGCCCTGCTTCTTCAACCAACTGCATCCCTAACTGCCCTGCTTCTTCTGT 2356
Db 2225 AACACATGGCCCTGCTTCTTCAACCAACTGCATCCCTAACTGCCCTGCTTCTTCTGT 2284
Qy 2357 CCTTCTCTCTTTTAATAAACCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCCTTC 2416
Db 2285 CCTTCTCTCTTTTAATAAACCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCCTTC 2344
Qy 2417 TGGTGGTAGTCCCACTTCTGCGATGCTCAATCCCTTCTCTACATCTTGTTCAAATCCTC 2476
Db 2345 TGGTGGTAGTCCCACTTCTGCGATGCTCAATCCCTTCTCTACATCTTGTTCAAATCCTC 2404
Qy 2477 ACTTTAAGGAGGATCTGGTAGCCCTGAGAAAGCAAACTAGCTCTGGACAAGATCAAAAC 2536
Db 2405 ACTTTAAGGAGGATCTGGTAGCCCTGAGAAAGCAAACTAGCTCTGGACAAGATCAAAAC 2464
Qy 2537 ACCCAAGCTTGATGCTCAATTAACCTCGATGATGCGAAAAACAGTCTCTGTGACTCAAACTC 2596
Db 2465 ACCCAAGCTTGATGCTCAATTAACCTCGATGATGCGAAAAACAGTCTCTGTGACTCAAACTC 2524
Qy 2597 AAGCCTTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCGCTCCAGTTCCGTGC 2656
Db 2525 AAGCCTTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCGCTCCAGTTCCGTGC 2584
Qy 2657 CATCACGAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCCAT 2716
Db 2585 CATCACGAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCCAT 2644
Qy 2717 GTCT 2720
Db 2645 GTCT 2648
```

Search completed: July 13, 2005, 03:40:20
Job time : 1625 secs

This Page Blank (uspto)

IF	0	0.5	1.0	2	AS/003	ups family protein

Qy

Qy

Db 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Qy 61 GGGGGCAGCTCTCCAGGCTGGTGTGTCTGAGGGGCTGCCCAACACATCTGTCATG 120
Db 21 GlyGlySerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Qy 121 GAGCCCAAGCAGAGATGTTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Qy 181 CTTTCCAACTCAGCGCTTCCACCTCCTACCTAGACCTCAGTATGAACAAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAATCCCTGCCAGCTCCGCTCTCCGAGGAGTTACCTCTTGGGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
Qy 301 GCTCTGACATACATCCCAAGGAGCATTCACCTGGCCCTTACAGTCTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGAAATATCAGCTAAGACACGCTACCCACAGAAGCTCTGCAGAAATTTGCCGAAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCTCGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATTCCTTGAGCACCTGTGCTGGATGACAACTGGTTACAGAAATCCCGCTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTAAAGATTATCGGATTCGAAGCCATGACCTTGGCCCTGGAACAAATAACACAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Qy 601 ATACCACTATGCTTTGAAACCTCTCCAGCTGTGTAGTTTACATCTCCATTAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuLeuHisLeuHisAsnAsn 220
Qy 661 AGATCCACTCCCTGGGAAGAATGCTTTCATGGGCTCCACAGCTAGAGACTTATAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTTAAATTAACAATAACCTTGTATGAAATCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAATAGGATTTATAGACAATATATCAGGTCGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CCTCTCTTATCAATATCAATCTTATGACAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTTACCTGAATAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTGATTTAACTGGAACCTGAAACCTGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCAAAACCGCTGCAATCAGTTACTTAATCTCCAAGTCTAGTCTCTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTATCCAGGTTTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1140

Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTTAAGACATTAATGAATCTACGAAATTAAGATTGACACTTTTCAGGAGTTCCTTAGGCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTCGAAACAAATTTGCTATTATTACCCCAATGATATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCCATCCCTAATAAAGCTGGACCTATCTGCCAACCTCTCTGCTGCTTTTTCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTAACTGACTTAATAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAACCTTCCAGAACTCAAGTTATAGAATGCTTATGCTTACCAAGTGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGACACACCTTCATAGAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCTGCTTGTGCTTGTAGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTTCCCGAGCCCTTTCAAAACCTGTGAACACCTGTGTGATGCTGCTGCATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATGGAGTGGACCATAGCAGTTCGGCACTTCTGTAATGCTTGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACTGTTAATTTGGGGTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGTCTACGGGAGTCTCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTCACGACATGCTGCTGGTGGGAGAAATGGGTTGGTTCCTGCTGCTGCT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTGTGCTTTCAGAAATCATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Qy 1981 GAGCGTGGTTCCTCTGTAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAATATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
Qy 2101 CTGGTGGCAGCAGATGATGGCTCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740

QY 2221 ACCATTGCTACCAAGCTCTACTGCAATTGTCAGAGGAGGAGACCTGGAGATATTGG 2280
Db 741 ThrileAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
QY 2281 GACTGCTCTATGTTAAACACATTCGCTCTGCTCTTCCACCAACTGCATCTTAACCTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTCTTGTCTCTCTCTCTTATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTTCTGGTGTAGTCCCATCTCTCTGCATGTCCTCAATCCCTTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTGTTCATCTCTCTTAAAGGAGCATCTGGTGAGCCCTGAGAAACCAACCTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACTCTGTATGTTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGATCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCATTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCGGTCCATCACCAGCTTATCCAGTACGTAGAGTGCATCTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907
RESULT 2
JG0193
G protein-coupled receptor FEX - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A:Title: Identification of a novel seven-transmembrane receptor with homology to glycoprotein 130
A:Reference number: JG0193; MUID:99121227; PMID:9920770
A:Accession: JG0193
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>
Alignment Scores:
Pred. No.: 2,15e-43 Length: 907
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.62% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x JG0193 (1-907)
QY 607 GACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAATAGATC 666
Db 203 AspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnArgIle 222
QY 667 CACTCCCTGGAAAGAAATGCTTTGATGGCTCCACAGCCCTAGACACTTTAGATTTAAAT 726
Db 223 HisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAspLeuAsn 242
QY 727 TACATAAACCTTGATGAATTCCTCCACTGCAATT 759
Db 243 TyrAsnAsnLeuAspGluPheProThrAlaIle 253

RESULT 3
AF2638
conserved hypothetical protein Atu0506 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF2638
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.; Wood, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF2638
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-195 <KUR>
A:Cross-references: UNIPROT:Q8UHZ5; GB:AB008688; PIDN:AAL41524.1; PID:g17738854; GSPDB:G17738854
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0506
A:Map position: circular chromosome
Alignment Scores:
Pred. No.: 6.31 Length: 195
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x AF2638 (1-195)
QY 73 CCCAGGTCGTGCTGCTGCTGAGGGGC 99
Db 145 ProArgSerGlyValLeuLeuArgGly 153
RESULT 4
G97420
hypothetical protein AGR_C_894 [imported] - Agrobacterium tumefaciens (strain C58, Cerezo)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97420
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209 <KUR>
A:Cross-references: UNIPROT:Q8UHZ5; GB:AB007869; PIDN:AAK86320.1; PID:g15155438; GSPDB:G15155438
C:Genetics:
A:Gene: AGR_C_894
A:Map position: circular chromosome
Alignment Scores:
Pred. No.: 6.25 Length: 209
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x G97420 (1-209)
QY 73 CCCAGGTCGTGCTGCTGCTGAGGGGC 99
Db 159 ProArgSerGlyValLeuLeuArgGly 167

```
RESULT 5
T22541
hypothetical protein F53B6.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T22541
R/White, S.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19578
A/Accession: T22541
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-342 <WIL>
A/Cross-references: UNIPROT:Q9XVW1; EMBL:Z81086; PIDN:CA803117.1; GSPDB:GN00019; CESP:FS
A/Experimental source: clone F53B6
C/Genetics:
A/Gene: CESP:F53B6.4
A/Map position: 1
A/Introns: 70/3; 212/3; 314/3

Alignment Scores:
Pred. No.: 5.85 Length: 342
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x T22541 (1-342)
Qy 1592 TCCTCAAGTCAAGCAGGAAATCTTCA 1566
Db 170 SerSerLySerSerArgLySerSer 178

RESULT 6
S06057
gene ND1 intron 3 protein 2 - Podospora anserina mitochondrion
C/Species: mitochondrion Podospora anserina
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: S06057
R/Cummings, D.J.; Domencico, J.M.; Michel, F.
Curr. Genet. 14, 253-264, 1988
A/Title: DNA sequence and organization of the mitochondrial ND1 gene from Podospora anse
A/Reference number: S06056; MUID:89063443; PMID:3197134
A/Accession: S06057
A/Molecule type: DNA
A/Residues: 1-580 <CUM>
A/Cross-references: UNIPROT:Q02714; EMBL:X13164
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SCC3
C/Keywords: mitochondrion

Alignment Scores:
Pred. No.: 5.46 Length: 580
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S06057 (1-580)
Qy 1588 GAGGAAGACCTGAAAGCCCTTCATCA 1614
Db 108 GluGluAspLeuLysAlaLeuHisSer 116

RESULT 7
A71400
probable disease resistance protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
A/Variety: Columbia
C/Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
```

```
C/Accession: A71400
R/Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giele
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech,
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ansc
C.; Chalwatzis, N.
A/Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A/Reference number: A71400; MUID:98121113; PMID:9461215
A/Accession: A71400
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-869 <BEV>
A/Cross-references: UNIPROT:O23253; GB:Z97335; NID:G2244747; PID:e326857; PID:G2244748
C/Genetics:
A/Map position: 4COP9-4G3845

Alignment Scores:
Pred. No.: 5.17 Length: 869
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x A71400 (1-869)
Qy 1261 TTCCCATCCCTAATAAAGCTGACCTA 1287
Db 334 LeuProSerLeuIleLysLeuAspLeu 342

RESULT 8
S77257
hypothetical protein s11265 - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S77257
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S77257
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1978 <KAN>
A/Cross-references: UNIPROT:P73551; EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BAAL17591
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Alignment Scores:
Pred. No.: 4.63 Length: 1978
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S77257 (1-1978)
Qy 997 CTGACTTTAACTGAGCAGATCTCA 1023
Db 526 LeuThrLeuThrGlyAlaGlnIleSer 534

RESULT 9
S07957
hypothetical protein P-plus - Escherichia coli
C/Species: Escherichia coli
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 08-Oct-1999
C/Accession: S07957
R/Plasterk, R.H.A.; van de Putte, P.
```

EMBO J. 4, 237-242, 1985
A:Title: The invertible P-DNA segment in the chromosome of *Escherichia coli*.
A:Reference number: S07185; MUID:85257443; PMID:3894006
A:Accession: S07957
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-145 <PLA>
A:Cross-references: EMBL:X01805; NID:942407; PIDN:CAA25947.1; PID:9581189
C:Genetics:
A:Map position: 25 min
A:Start codon: GTG

Alignment Scores:
Pred. No.: 70.8 Length: 145
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x S07957 (1-145)
Qy 988 GGTTCAGTTCACAGTTAAATCAG 965
Db 128 GlyLeuGlnPheGlnLeuAsnGln 135

RESULT 10
AB1265
hypothetical protein lmol522 [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1265
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <GLA>
A:Cross-references: UNIPROT:Q93RD9; GB:NC_003210; PIDN:CAC99600.1; PID:g16410951; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmol522
C:Superfamily: conserved hypothetical protein HI0670

Alignment Scores:
Pred. No.: 70.5 Length: 150
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x AB1265 (1-150)
Qy 2136 GCAGAGGGGAGGGCCATCTT 2113
Db 67 AlaGluArgGlyAlaIleLeu 74

RESULT 11
AD1627
hypothetical protein homolog lin157 [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1627
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <GLA>
A:Cross-references: UNIPROT:Q92BJ1; GB:AL592022; PIDN:CAC96788.1; PID:g16414044; GSPDB:
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1557
C:Superfamily: conserved hypothetical protein HI0670

Alignment Scores:
Pred. No.: 70.5 Length: 150
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x AD1627 (1-150)
Qy 2136 GCAGAGGGGAGGGCCATCTT 2113
Db 67 AlaGluArgGlyAlaIleLeu 74

RESULT 12
T31777
hypothetical protein F38H12.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31777
R:Wessner, J.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of *C. elegans* cosmid F38H12.
A:Reference number: Z21084
A:Accession: T31777
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-158 <WOE>
A:Cross-references: UNIPROT:O16345; EMBL:AF016436; PIDN:AAC25897.1; GSPDB:GN00023; CESP:
A:Experimental source: strain Bristol N2; clone F38H12
C:Genetics:
A:Gene: CESP:F38H12.5
A:Map position: 5
A:Introns: 27/1; 38/3; 107/3

Alignment Scores:
Pred. No.: 70 Length: 158
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x T31777 (1-158)
Qy 1226 TTGTTCCAGCCAAATTCAGCGAT 1203
Db 151 LeuPheGlnAlaLysPheSerAsp 158

RESULT 13
F97440
hypothetical protein AGR_C_1201 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
C:Species: *Agrobacterium tumefaciens*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97440
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wolam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.;
Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: F97440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: UNIPROT:Q8U5I8; GB:AB007869; PIDN:AAK86479.1; PID:g15155627; GSPDB:G
A;Gene: AGR_C_1201
A;Map position: circular chromosome

Alignment Scores:
Pred. No.: 69.8 Length: 162
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x F97440 (1-162)

Qy 2068 CTGCTGGCCTTGACCATGGCGCA 2091

Db 22 LeuLeuAlaLeuThrMetAlaAla 29

RESULT 14

Dps family protein [imported] - Caulobacter crescentus
A;Accession: A87605
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.D.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-178 <STO>
A;Cross-references: UNIPROT:Q9A4G1; GB:AB005673; NID:g13424489; PIDN:AAK24837.1; GSPDB:G
C;Genetics:
A;Gene: CC2873

Alignment Scores:
Pred. No.: 68.9 Length: 178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x A87605 (1-178)

Qy 613 GCCTTGGAACTCTCCAGCTTG 636

Db 104 AlaPheGlyAsnLeuSerSerLeu 111

RESULT 15

S75434
hypothetical protein c05017 - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432; PMID:8899719

A;Accession: S75434
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-179 <SEN>

A;Cross-references: UNIPROT:P96006; EMBL:Y08257; NID:g1707772; PID:e283869; PID:g1707840
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Genetics:
A;Start codon: GTG

Alignment Scores:
Pred. No.: 68.8 Length: 179
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S75434 (1-179)

Qy 2035 TCTAGCCTGAAAGTAATCATTTTG 2058

Db 14 SerSerLeuLysValIleIleLeu 21

Search completed: July 13, 2005, 03:58:32
Job time : 83.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 20:27:05 ; Search time 9588 Seconds
(without alignments)
12073.483 Million cell updates/sec

Title: US-10-751-736-21
Perfect score: 2724
Sequence: 1 atgacacctccggctgg.....catttgtcccatgtctctaa 2724

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2296	84.3	2724	9 AY411732	Homo sapi
2	951	34.9	1060	1 AL541959	AL541959
3	706	25.9	800	4 BG196708	RST15934
4	669	24.6	1104	1 AL530798	AL530798
5	620	22.8	671	4 BW795259	K-EST0077
6	538	19.8	590	6 CD617354	CD617354
7	529	19.4	873	5 BX403219	BX403219
8	475	17.4	475	5 BX107244	BX107244
9	472	17.3	472	1 AA460529	AA460529
10	418	15.3	556	5 BQ323949	BQ323949
11	408	15.0	2593	9 AY411733	Pan trogl
12	369	13.5	374	1 A1367723	A1367723
13	354	13.0	358	1 A1699236	A1699236
14	345	12.7	494	1 AA424098	AA424098
15	341	12.5	539	1 A1697103	A1697103
16	275	10.1	539	5 BX499910	BX499910
17	252	9.3	276	6 CD617347	CD617347
18	252	9.3	276	6 CD617349	CD617349
19	252	9.3	276	6 CD617353	CD617353
20	251	9.2	276	6 CD617346	CD617346
21	251	9.2	276	6 CD617348	CD617348
22	246	9.0	271	6 CD617352	CD617352
23	245	9.0	268	6 CD617342	CD617342
24	235	8.6	250	6 CD617356	CD617356

c 25	222	8.1	222	6	CD617357	CD617357	55041447J
c 26	222	8.1	222	6	CD617358	CD617358	55041471H
c 27	219	8.0	273	6	CD617343	CD617343	55013050J
c 28	201	7.4	275	6	CD617351	CD617351	55013056J
c 29	200	7.3	275	6	CD617350	CD617350	55013056H
c 30	193	7.1	330	5	BX357656	BX357656	BX357656
c 31	176	6.5	224	6	CD617359	CD617359	55041471J
c 32	176	6.5	432	8	AQ076504	AQ076504	CIT-HSP-2
c 33	169	6.2	275	6	CD617344	CD617344	55013051H
c 34	146	5.4	492	6	CD617355	CD617355	55041415J
c 35	139	5.1	397	1	A1539105	A1539105	tp78b01.x
c 36	125	4.6	542	1	A1962439	A1962439	wg52c12.x
c 37	119	4.4	396	1	AL113316	AL113316	DKF2p761C
c 38	112	4.1	940	7	CK449409	CK449409	892976 MA
c 39	96	3.5	558	2	BE929414	BE929414	NRO-GN002
c 40	94	3.5	215	7	CO933257	CO933257	AGENCOURT
c 41	94	3.5	280	6	CD617345	CD617345	55013051J
c 42	90	3.3	391	8	B53762	B53762	CIT-HSP-201
c 43	87	3.2	324	1	A1206192	A1206192	qr2xd02.x
c 44	87	3.2	669	9	CE133614	CE133614	tigr-ges-
c 45	83	3.0	634	7	CF177761	CF177761	806271 MA

ALIGNMENTS

RESULT 1	AY411732	2724 bp	DNA	linear	GSS 16-DEC-2003
LOCUS	AY411732	Homo sapiens GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY411732	Genomic survey sequence.			
ACCESSION	AY411732	GI:39767700			
VERSION	AY411732.1	GI:39767700			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2724)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..2724				
gene	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	<1..>2724				
	/gene="GPR49"				
	/locus_tag="HCM4322"				
ORIGIN					
Query Match	84.3%;	Score 2296;	DB 9;	Length 2724;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	429	GGGTCTGGAGCTAACACATCAGCTATGTGCCCCAGCTGTTTCAGTCGCTGCATTC	488		

Db 429 GCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTTCAGTGCCTGCATTC 488
Qy 489 CCTGAGGACCTGTGGTGGATGACAAATGGCTTTAAAGAAATCCCGCTCCAGGCTTTTAG 548
Db 489 CCTGAGGACCTGTGGTGGATGACAAATGGCTTTAAAGAAATCCCGCTCCAGGCTTTTAG 548
Qy 549 AAGTTTATCGGATGCAAGCCATGACCTTGGCCCTGAAACAAATACACACATACACGA 608
Db 549 AAGTTTATCGGATGCAAGCCATGACCTTGGCCCTGAAACAAATACACACATACACGA 608
Qy 609 CTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATACCTCCATAACAAATAGAAATCCA 668
Db 609 CTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATACCTCCATAACAAATAGAAATCCA 668
Qy 669 CTCCCTGGGAAGAAATGCTTTGTATGGCTCCACAGCTAGAGACTTTAGATTTAAATTA 728
Db 669 CTCCCTGGGAAGAAATGCTTTGTATGGCTCCACAGCTAGAGACTTTAGATTTAAATTA 728
Qy 729 CAATAACCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAAACCTTAAAGAACTAGG 788
Db 729 CAATAACCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAAACCTTAAAGAACTAGG 788
Qy 789 ATTTTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAACCTTCTCT 848
Db 789 ATTTTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAACCTTCTCT 848
Qy 849 TATTACAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCTTTTCAACA 908
Db 849 TATTACAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCTTTTCAACA 908
Qy 909 TTTTACCTGAACTAGAACACTGACTCTGAATGGTGGCTCCACAATATCAATTTCCCTGA 968
Db 909 TTTTACCTGAACTAGAACACTGACTCTGAATGGTGGCTCCACAATATCAATTTCCCTGA 968
Qy 969 TTTTAACTGGAACTGCAAAACCTGGAGAGTCTGACCTTTAACTGGAGCACAGATCTCATCTCT 1028
Db 969 TTTTAACTGGAACTGCAAAACCTGGAGAGTCTGACCTTTAACTGGAGCACAGATCTCATCTCT 1028
Qy 1029 TCCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCAAAGTGTAGATCTGCTTTACAACCT 1088
Db 1029 TCCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCAAAGTGTAGATCTGCTTTACAACCT 1088
Qy 1089 ATTAGAAGATTTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGACCTTAAGACA 1148
Db 1089 ATTAGAAGATTTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGACCTTAAGACA 1148
Qy 1149 TAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCCGATCGCT 1208
Db 1149 TAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCCGATCGCT 1208
Qy 1209 GAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTGCCATC 1268
Db 1209 GAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTGCCATC 1268
Qy 1269 CCTAATAAGCTGACCTATGCTCCAACTCTGCTGCTTTTCCCTATTAACCTGGTTACA 1328
Db 1269 CCTAATAAGCTGACCTATGCTCCAACTCTGCTGCTTTTCCCTATTAACCTGGTTACA 1328
Qy 1329 TGGTTTAACTCACATTTAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCCGATCGCT 1388
Db 1329 TGGTTTAACTCACATTTAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCCGATCGCT 1388
Qy 1389 AAACTTTCCGAACTCAAGGTTATAGAAATGCTTTATGCTTTACAGAGCTTTGATATCATCTGA 1448
Db 1389 AAACTTTCCGAACTCAAGGTTATAGAAATGCTTTATGCTTTTACCAGTGTGCTGCAATTTGG 1448
Qy 1449 AGTGTGTGAGAACTGCTATAGATTTCTTAATCAATGGAATTAAGGTGACACACAGTAT 1508
Db 1449 AGTGTGTGAGAACTGCTATAGATTTCTTAATCAATGGAATTAAGGTGACACACAGTAT 1508
Qy 1509 GGACGACCTTCATAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGTGACCTTGA 1568
Db 1509 GGACGACCTTCATAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGTGACCTTGA 1568

Qy 1569 AGATTTTCTGCTTGACTTTGAGGAAGACCTTGAAGCCCTTCATTTCACTGACGTGTTTCAACC 1628
Db 1569 AGATTTTCTGCTTGACTTTTGAAGGAAGACCTTGAAGCCCTTCATTTCACTGACGTGTTTCAACC 1628
Qy 1629 TTCCCCAGGGCCCTTCAAAACCTGTGAACACCTGTCTGTATGGCTGGCTGATCAAGATTGG 1688
Db 1629 TTCCCCAGGGCCCTTCAAAACCTGTGAACACCTGTCTGTATGGCTGGCTGATCAAGATTGG 1688
Qy 1689 AGTGTGGACCATAGCAGTTCTTGGCAGTCTACTTGTAAATGCTTTTGGTGACTTTCAACAGTTTT 1748
Db 1689 AGTGTGGACCATAGCAGTTCTTGGCAGTCTACTTGTAAATGCTTTTGGTGACTTTCAACAGTTTT 1748
Qy 1749 CAGATCCCTCTGTGTACATTTTCCCCCATTAACCTGTTAAATGGGGTCAATCCGACAGTGAA 1808
Db 1749 CAGATCCCTCTGTGTACATTTTCCCCCATTAACCTGTTAAATGGGGTCAATCCGACAGTGAA 1808
Qy 1809 CATGCTCAGCGGAGTCTCCAGTGCCTGTGGTGTGGTGTGATGCTTCACTTTTGGCAG 1868
Db 1809 CATGCTCAGCGGAGTCTCCAGTGCCTGTGGTGTGGTGTGATGCTTCACTTTTGGCAG 1868
Qy 1869 CTTTGCACGACATGCTGTGGGAGAAATGGGGTTGGTGGCAATGTCAATTTGGTTTTTT 1928
Db 1869 CTTTGCACGACATGCTGTGGGAGAAATGGGGTTGGTGGCAATGTCAATTTGGTTTTTT 1928
Qy 1929 GTCCATTTTGTCTCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGG 1988
Db 1929 GTCCATTTTGTCTCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGG 1988
Qy 1989 GTTCTCTGTGAAATATTCTGCAAAATTTGAAAAGAAAGCTCCATTTTCTAGCCCTGAAAGT 2048
Db 1989 GTTCTCTGTGAAATATTCTGCAAAATTTGAAAAGAAAGCTCCATTTTCTAGCCCTGAAAGT 2048
Qy 2049 AATCATTTTGTCTGTGCTTGTGCTTGTGACCATGGCCGAGTTCCCTGCTGGGTGG 2108
Db 2049 AATCATTTTGTCTGTGCTTGTGCTTGTGACCATGGCCGAGTTCCCTGCTGGGTGG 2108
Qy 2109 CAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGCCCGACCACT 2168
Db 2109 CAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGCCCGACCACT 2168
Qy 2169 GGGCTACATGGTGTGCTCTCATCTTGTCAAATCCCTTGTCTTCTCATGATGACCAATGC 2228
Db 2169 GGGCTACATGGTGTGCTCTCATCTTGTCAAATCCCTTGTCTTCTCATGATGACCAATGC 2228
Qy 2229 CTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTC 2288
Db 2229 CTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTC 2288
Qy 2289 TATGGTAAACACATTTGGCGCTTGTCTTCAACCACTGCATCTTAACTGCCCTGTGGC 2348
Db 2289 TATGGTAAACACATTTGGCGCTTGTCTTCAACCACTGCATCTTAACTGCCCTGTGGC 2348
Qy 2349 TTTCTTGTCTCTCTCTTTTAAATAAACCTTATATATATATATATATATATATATATATAT 2408
Db 2349 TTTCTTGTCTCTCTCTTTTAAATAAACCTTATATATATATATATATATATATATATATAT 2408
Qy 2409 TATCTTCTGGTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTTACATCTTGT 2468
Db 2409 TATCTTCTGGTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTTACATCTTGT 2468
Qy 2469 CAATCTCAGCTTTAAAGGAGATCTGGTGGAGCCCTGAGAAAGCAACCTTACCTCTGGACAAG 2528
Db 2469 CAATCTCAGCTTTAAAGGAGATCTGGTGGAGCCCTGAGAAAGCAACCTTACCTCTGGACAAG 2528
Qy 2529 ATCAAAACACCCAGCTTGTATGATTAATTAATCTCTGATGATGTCGAAAAACAGTCTGTGA 2588
Db 2529 ATCAAAACACCCAGCTTGTATGATTAATTAATCTCTGATGATGTCGAAAAACAGTCTGTGA 2588
Qy 2589 CTCAACTCAAGCCCTTGGTAACTTTTACAGCTCAGCATCACTTATGACCTGCCCTCCAG 2648
Db 2589 CTCAACTCAAGCCCTTGGTAACTTTTACAGCTCAGCATCACTTATGACCTGCCCTCCAG 2648

3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherays.com
High quality sequence stop: 481.

FEATURES
source

ORIGIN

Query Match	25.9%;	Score 706;	DB 4;	Length 800;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 706;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1210	AAATTTGCTCTGGACAAAAATTTGCTATTATTATTCACCCCAATGCATTTTCCACTTTGGCCATCC	1269	
Db	1	AAATTTGCTCTGGACAAAAATTTGCTATTATTATTCACCCCAATGCATTTTCCACTTTGGCCATCC	60	
Qy	1270	CTAATAAGCTGGACCTATCGTCCAACTCCTGTCGTCTTTTCCCTATAAAGCTGGGTACAT	1329	
Db	61	CTAATAAGCTGGACCTATCGTCCAACTCCTGTCGTCTTTTCCCTATAAAGCTGGGTACAT	120	
Qy	1330	GGTTTAACTCATTTAAATTTAAACAGGAAATCATGCCTTACAGACTTGATCATCTGAA	1389	
Db	121	GGTTTAACTCATTTAAATTTAAACAGGAAATCATGCCTTACAGACTTGATCATCTGAA	180	
Qy	1390	AACTTTCCAGAACTCAAGGTTTATAGAAATGCCTTATGCTTACCAAGTCGTGCAATTTGGA	1449	
Db	181	AACTTTCCAGAACTCAAGGTTTATAGAAATGCCTTATGCTTACCAAGTCGTGCAATTTGGA	240	
Qy	1450	GTGTGTGAGAAATGCTTAAAGATTTCTAAATCAATTGGAAATAAGGTGACAAACAGCAGTATG	1509	
Db	241	GTGTGTGAGAAATGCTTAAAGATTTCTAAATCAATTGGAAATAAGGTGACAAACAGCAGTATG	300	
Qy	1510	GACGACCTTCANTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACCTTGAA	1569	
Db	301	GACGACCTTCANTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACCTTGAA	360	
Qy	1570	GATTTTCCTGCTTGACTTTGAGGAGAGACCTGAAAGCCCTTCATTTACAGTGCAGTGTTCACCT	1629	
Db	361	GATTTTCCTGCTTGACTTTGAGGAGAGACCTGAAAGCCCTTCATTTACAGTGCAGTGTTCACCT	420	
Qy	1630	TCCCCAGGCCCTTTCAAACCCCTGTGGAACACCTGCTTGATGGCTGGCTGATCAGAAATTTGGA	1689	
Db	421	TCCCCAGGCCCTTTCAAACCCCTGTGGAACACCTGCTTGATGGCTGGCTGATCAGAAATTTGGA	480	
Qy	1690	GTGTGGACCATAGCAGTTCCTGGCACTTACTGTGTAATGCTTTGTGTGACTTCAACAGATTTTC	1749	
Db	481	GTGTGGACCATAGCAGTTCCTGGCACTTACTGTGTAATGCTTTGTGTGACTTCAACAGATTTTC	540	
Qy	1750	AGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGTTCATGCGACGATGAAC	1809	
Db	541	AGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATTTGGGTTCATGCGACGATGAAC	600	
Qy	1810	ATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGTGATGGTTCACCTTTTGGCAGC	1869	
Db	601	ATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGTGATGGTTCACCTTTTGGCAGC	660	
Qy	1870	TTTGACGACATGCTGCTGGTGGGAAATGGGGTTGGTGGCCATG	1915	
Db	661	TTTGACGACATGCTGCTGGTGGGAAATGGGGTTGGTGGCCATG	706	

RESULT	AL530798
LOCUS	AL530798 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CSDD008YF22 5-PRIME, mRNA sequence.
DEFINITION	linear EST 24-MAR-2004 1104 bp mrna
ACCESSION	AL530798
VERSION	AL530798.3 GI:45705748
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Verteabrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1104) Li,W.B., Gruber,C., Jessee,J. and Polayes,D. Full-length cdNA libraries and normalization Unpublished (2001) On Feb 13, 2001 this sequence version replaced gi:31068631. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr The first strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7021.f For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CSDD008DC110P1&c=7021.f .
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

```
Db 361 TCATGCCTTACAGAGCTTGATATCATCTGAAACCTTCCAGAACTCTCAGAACTATAGAAAT 420
Qy 1419 GCCT-TATGCTTACCAGTCTGTGCAATTTGGAGTGTGAGAAATGCCTATAGATTTCTTA 1477
Db 421 GCCTATATGCTTACCAGTCTGTGCAATTTGGAGTGTGAGAAATGCCTATAGATTTCTTA 480
Qy 1478 ATCAATGGAATTAAGGTGACACACAGCATGATGGACGCTTCTTAAAGAAAGTCTGGNA 1537
Db 481 ATCAATGGAATTAAGGTGACACACAGCATGATGGACGCTTCTTAAAGAAAGTCTGGNA 540
Qy 1538 TGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGACCTTTGAGGAGAC 1597
Db 541 TGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGACCTTTGAGGAGAC 600
Qy 1598 TGAAGGCCCTTCATTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAC 1657
Db 601 TGAAGGCCCTTCATTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAC 660
Qy 1658 ACCTGCTTGATGGCTGGTGTGATCAGAAATTTGGAGTGTGGACCATAGCAGTCTGGCACTTA 1717
Db 661 ACCTGCTTGATGGCTGGTGTGATCAGAAATTTGGAGTGTGGACCATAGCAGTCTGGCACTTA 720
Qy 1718 CTGTGTAATGCTTTGTGTGACTCAACAGTCTTCCAGATCCCTCTGTACATTTTCCCCCATTA 1777
Db 721 CTGTGTAATGCTTTGTGTGACTCAACAGTCTTCCAGATCCCTCTGTACATTTTCCCCCATTA 780
Qy 1778 AACTGTTAATTTGGGTCATCGCAGCATGAACATGCTCACGGGAGTCTCCAGTGCCTGTC 1837
Db 781 AACTGTTAATTTGGGTCATCGCAGCATGAACATGCTCACGGGAGTCTCCAGTGCCTGTC 840
Qy 1838 TGGCTGGTGTGATGCGTTCATCTTTGGCAGCTTTTGCAGCATGCTGGTGGGAGA 1897
Db 841 TGGCTGGTGTGATGCGTTCATCTTTTGGCAGCTTTTGCAGCATGCTGGTGGGAGA 900
Qy 1898 ATGGGCTTGGTGGTGCATGCTATGCTTTTGTGCTTTCCTATTTTTC 1940
Db 901 ATGGGCTTGGTGGTGCATGCTATGCTTTTGTGCTTTCCTATTTTTC 943

RESULT 5
BM795259 671 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0077025 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-31-C01 5',
DEFINITION mRNA sequence.
ACCESSION BM795259
VERSION BM795259.1 GI:19143491
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 671)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,O.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 31 row: C column: 01
High quality sequence stop: 671.
Location/Qualifiers
1. .671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-31-C01"
/sex="F"

FEATURES
source
```

```
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ189pl; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 22.8%; Score 620; DB 4; Length 671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1775 TTAAACTCTTAATTTGGGTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGC 1834
Db 1 TTAAACTCTTAATTTGGGTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGC 60
Qy 1835 TGCTGGCTGGTGTGGATCGTTCATCTTTTGGCAGCTTTTGCACGACATGGTGGTGG 1894
Db 61 TGCTGGCTGGTGTGGATCGTTCATCTTTTGGCAGCTTTTGCACGACATGGTGGTGG 120
Qy 1895 AGAATGGGCTGGTGGTGGATCGTTCATCTTTTGGTCCATTTTGTCTTCAGATCATCTG 1954
Db 121 AGAATGGGCTGGTGGTGGATCGTTCATCTTTTGTCTTCATTTTGTCTTCAGATCATCTG 180
Qy 1955 TTTTCTCTCTTACTCTTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAAATATTCTGCAAAAT 2014
Db 181 TTTTCTCTCTTACTCTTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAAATATTCTGCAAAAT 240
Qy 2015 TTGAAACGAAAGCTCCATTTTCTAGCTTGAAGTAATCATTTTGTGCTGTGCTGCTG 2074
Db 241 TTGAAACGAAAGCTCCATTTTCTAGCTTGAAGTAATCATTTTGTGCTGTGCTGCTGCTG 300
Qy 2075 CCTTGACCATGGCCGCGAGTTCCTCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCT 2134
Db 301 CCTTGACCATGGCCGCGAGTTCCTCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCT 360
Qy 2135 GCCTGCCCTTTGCCCTTTTGGGAGCCGACCATGGGTACATGCTGCTCTCATCTTTC 2194
Db 361 GCCTGCCCTTTGCCCTTTTGGGAGCCGACCATGGGTACATGCTGCTCTCATCTTTC 420
Qy 2195 TCAATTCCTTTTGTCTCTCATGATGACCATTTGCCATACCAAGCTCTACTGCAATTTGG 2254
Db 421 TCAATTCCTTTTGTCTCTCATGATGACCATTTGCCATACCAAGCTCTACTGCAATTTGG 480
Qy 2255 ACAAGGAGAGCTGGAGAAATATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTTGGC 2314
Db 481 ACAAGGAGAGCTGGAGAAATATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTTGGC 540
Qy 2315 TCTTCACCAACTGCATCTTAACTGCCCTGGGCTTCTGTGCTCTCTCTCTCTTAAATAA 2374
Db 541 TCTTCACCAACTGCATCTTAACTGCCCTGGGCTTCTGTGCTCTCTCTCTCTTAAATAA 600
Qy 2375 ACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCTCTTCTGGTGGTAGTCCACTTC 2434
Db 601 ACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCTCTTCTGGTGGTAGTCCACTTC 660
Qy 2435 CTGATGCTC 2445
|||||
```

```
Db      661 CTGCATGCTC 671

RESULT 6
CD617354
LOCUS   CD617354          590 bp      mRNA      linear      EST 12-JAN-2004
DEFINITION 55041415H1 PFP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD617354
VERSION   CD617354.1 GI:40265619
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS  Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE    Circular rapid amplification of cDNA ends for high-throughput
          extension cloning of partial genes
JOURNAL  Genomics 84 (1), 205-210 (2004)
COMMENT  Contact: Fu GK
          Incyte Genomics, Inc.
          3160 Porter Dr., Palo Alto, CA 94304, USA
          Tel: 6508454102
          Email: gfu@incyte.com.
          Location/Qualifiers
            1..590
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="FLP"
              /note="Vector: pDrive Cloning Vector"

FEATURES             source
source               1..590
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone_lib="FLP"
                    /note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match      19.8%; Score 538; DB 6; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.8e-284;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 GCGGGAACGCTCGACATATCCCAAGGGAGCATTCAGTGGCTTTACAGTCTTAAA 351
Db 16 GCGGGAACGCTCGACATATTCACATATTCACAGGGAGCATTCAGTGGCTTTACAGTCTTAAA 75
Qy 352 GTTCTTATGCTGCAGAAATATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTG 411
Db 76 GTTCTTATGCTGCAGAAATATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTG 135
Qy 412 CGAAGCCTTCATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCAGCTGT 471
Db 136 CGAAGCCTTCATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCAGCTGT 195
Qy 472 TTCAGTGGCCTGCATTCCTCGAGGCACCTGTGGCTGGATGACATGCGTTAAACAGAAATC 531
Db 196 TTCAGTGGCCTGCATTCCTCGAGGCACCTGTGGCTGGATGACATGCGTTAAACAGAAATC 255
Qy 532 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATFTGCAAGCCATGACCTTTGGCCCTGAACAAA 591
Db 256 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATFTGCAAGCCATGACCTTTGGCCCTGAACAAA 315
Qy 592 ATACACACATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTTACATCTC 651
Db 316 ATACACACATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTTACATCTC 375
Qy 652 CATAACATGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAG 711
Db 376 CATACATAGATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAG 435
Qy 712 ACTTTAGATTTAAATTACAATAACCTTTGATGAATTCGCCATTCGCAATTAGGACACTCTCC 771
Db 436 ACTTTAGATTTAAATTACAATAACCTTTGATGAATTCGCCATTCGCAATTAGGACACTCTCC 495
Qy 772 AACCTTAAAGNACTAGGATTCATAGCAACATATCAGGTGCGATACCTTGAAAGCAT 829
Db 496 AACCTTAAAGNACTAGGATTTTCATAGCAACATATCAGGTGCGATACCTTGAAAGCAT 553
```

```
RESULT 7
BX403219
LOCUS   BX403219          873 bp      mRNA      linear      EST 29-APR-2004
DEFINITION BX403219 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
          clone CSODI029YG09 5-PRIME, mRNA sequence.
ACCESSION BX403219
VERSION   BX403219.2 GI:46875014
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 873)
AUTHORS  Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE    Full-length cDNA libraries and normalization
JOURNAL  Unpublished (2001)
COMMENT  On May 13, 2003 this sequence version replaced gi:30619059.
          Contact: Genoscope
          Genoscope - Centre National de Sequencage
          2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen. This sequence belongs to sequence cluster
          10101.f
          For more information about this cluster, see
          http://www.genoscope.cns.fr/cdna?s=CS2AAW3ZE04_AW9ZA03_l&c=10101.f.
          Location/Qualifiers
            1..873
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CSODI029YG09"
              /tissue_type="PLACENTA COT 25-NORMALIZED"
              /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
              /note="1st strand cDNA was primed with a NotI-oligo(dT)
              primer. Five prime end enriched, double-strand cDNA was
              digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.
              Library was normalized."

FEATURES             source
source               1..873
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CSODI029YG09"
                    /tissue_type="PLACENTA COT 25-NORMALIZED"
                    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                    /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.
                    Library was normalized."

ORIGIN
Query Match      19.4%; Score 529; DB 5; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.7e-279;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCACACCTCCCGGCTCGGTGTGCTCTCTGCTTGCCTGTGCTGCTGCTGCTGCGAC 60
Db 138 ATGCACACCTCCCGGCTCGGTGTGCTCTCTGCTTGCCTGTGCTGCTGCTGCGAC 197
Qy 61 GGGGGCAGCTCTCCAGGTCTGGTGTGCTGAGGGGCTGCCCCACACACTGTCATTC 120
Db 198 GGGGGCAGCTCTCCAGGTCTGGTGTGCTGAGGGGCTGCCCCACACACTGTCATTC 257
Qy 121 GAGCCGACGCGCAGGATGTTGCTCAGGTGACCTGCTCCGACCTGCGGGCTCTCGAGCTG 180
Db 258 GAGCCGACGCGCAGGATGTTGCTCAGGTGACCTGCTCCGACCTGCGGGCTCTCGAGCTG 317
Qy 181 CCTTCCAACTCAGCGCTTTCACCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
Db 318 CCTTCCAACTCAGCGCTTTCACCTCTACCTCCTA CTTAGACCTCAGTATGAACAACATCAGTCAG 377
Qy 241 CTGCTCCCGAATCCCTGCGCCAGTCTCCGCTTCTCTGAGGAGTTACCTCTTGGGGGAAC 300
Db 378 CTGCTCCCGAATCCCTGCGCCAGTCTCCGCTTCTCTGAGGAGTTACCTCTTGGGGGAAC 437
Qy 301 GCTCTGACATACATTCCTCCAGGGAGGAGTTCACCTGGCTTTTACAGTCTTAAAGTTCTTATG 360
Db 438 GCTCTGACATACATTCCTCCAGGGAGGAGTTCACCTGGCTTTTACAGTCTTAAAGTTCTTATG 497
Qy 361 CTGCAGAAATACAGCTTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
```

```

1755 CCCTCTGTACATTTCCCCCATTAACCTTTAAATTTGGGTTCATTCGACAGTGAACATGCT 1814
61 CCCTCTGTACATTTCCCCCATTAACCTTTAAATTTGGGTTCATTCGACAGTGAACATGCT 120
1815 CACGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1874
121 CACGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
1875 ACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1934
181 ACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
1935 TTTTGTCTCAGAAATCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1994
241 TTTTGTCTCAGAAATCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
1995 TGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCATTTCTAGCTGGAAGTAAATCAT 2054
301 TGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCATTTCTAGCTGGAAGTAAATCAT 360
2055 TTTGTCTCTGCTCCCTGCTGGCTTGACCATGGCCGAGTTCCTCTCTCTCTCTCTCTCTCTCTCT 2114
361 TTTGTCTCTGCTCCCTGCTGGCTTGACCATGGCCGAGTTCCTCTCTCTCTCTCTCTCTCTCTCT 420
2115 GTATGGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2169
421 GTATGGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 475

RESULT 9
AA460529 472 bp mRNA linear EST 09-JUN-1997
LOCUS zx68c09.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
DEFINITION IMAGE:796624 5', similar to WP:CS0H2.1 CE05479 STEROID HORMONE
RECEPTOR ;, mRNA sequence.
ACCESSION AA460529
VERSION AA460529.1 GI:2185649
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gaisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilton RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LILNL; contact the
IMAGE Consortium (info@image.lilnl.gov) for further information.
Seq primer: -28ml3 rev2 EF from Amersham.
Location/Qualifiers
1. 472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:6040989"
/db_xref="taxon:9606"
/clone="IMAGE:796624"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8 9w"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Patina Bonaldo."

FEATURES
source
Query Match 17.4%; Score 475; DB 5; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e-249;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1695 GACCATAGCAGTCTGGCAGCTTACTTGTAAAGCTTTGGTGTGACTTCAACAGTTTTCAGATC 1754
DB 1 GACCATAGCAGTCTGGCAGCTTACTTGTAAAGCTTTGGTGTGACTTCAACAGTTTTCAGATC 60

```

```

498 GTGAGAAATAATCAGCTAAGACACAGTACCACAGAGCTCTGCAGAAATTTGCGAGCCTT 557
421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCGCCCAAGCTGTTTCAGTGGC 480
558 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCGCCCAAGCTGTTTCAGTGGC 617
481 CTGCATTCCTGAGCAGCCTGTGGCTGGATGAGCAATGCGTTAAACAGAAA 529
618 CTGCATTCCTGAGCAGCCTGTGGCTGGATGAGCAATGCGTTAAACAGAAA 666

RESULT 8
BX107244 475 bp mRNA linear EST 06-FEB-2003
LOCUS BX107244 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
DEFINITION IMAGE:796624 5', similar to WP:CS0H2.1 CE05479 STEROID HORMONE
RECEPTOR ;, mRNA sequence.
ACCESSION BX107244
VERSION BX107244.1 GI:27847308
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE:998N171964.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCCACAGGAACAGCTATGAC.
Location/Qualifiers
1. 475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998N171964 ; IMAGE:796624"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8 9w"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Patina Bonaldo."

ORIGIN
Query Match 17.4%; Score 475; DB 5; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.6e-249;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1695 GACCATAGCAGTCTGGCAGCTTACTTGTAAAGCTTTGGTGTGACTTCAACAGTTTTCAGATC 1754
DB 1 GACCATAGCAGTCTGGCAGCTTACTTGTAAAGCTTTGGTGTGACTTCAACAGTTTTCAGATC 60

```


Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1 (bases 1 to 2593)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2593)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2593
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2593
/gene="GPR49"
/locus_tag="HCM4322"
ORIGIN
Query Match 15.0%; Score 408; DB 9; Length 2593;
Best Local Similarity 99.1%; Pred. No. 1.8e-212;
Matches 758; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1960 CTGCTTACTCGGACGCTGGAGCGTGGTCTCTGTAATATTCGAAATTTGCA 2019
Db 1829 CTGCTTACTCGGACGCTGGAGCGTGGTCTCTGTAATATTCGAAATTTGCA 1888
Qy 2020 ACGAAGCTCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGTGCGCTGTGCGCTTG 2079
Db 1889 ACGAAGCTCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGTGCGCTGTGCGCTTG 1948
Qy 2080 ACCATGGCGGAGTTCCCTGTGGTGGGACGAGTAATGGCGCTCCCTCTCTGCGCTG 2139
Db 1949 ACCATGGCGGAGTTCCCTGTGGTGGGACGAGTAATGGCGCTCCCTCTCTGCGCTG 2008
Qy 2140 CTTTTCCTTTTGGGAGCCGACGACCATGGCTACATGGTGGCTCTATCTTGTCTCAAT 2199
Db 2009 CTTTTCCTTTTGGGAGCCGACGACCATGGCTACATGGTGGCTCTATCTTGTCTCAAT 2068
Qy 2200 TCCCTTTGCTCTCATGATGACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAAG 2259
Db 2069 TCCCTTTGCTCTCATGATGACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAAG 2128
Qy 2260 GGAGACCTGGAGAAATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTC 2319
Db 2129 GGAGACCTGGAGAAATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTC 2188
Qy 2320 ACCAATCGATCTTAAACGCGCTGGCTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2379
Db 2189 ACCAATCGATCTTAAACGCGCTGGCTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2248
Qy 2380 ACAATTTATCAGTCTGAAGTAATTAAGTTTATCTCTGCTGGTGGTGGTGGTGGTGGTGG 2439
Db 2249 ACAATTTATCAGTCTGAAGTAATTAAGTTTATCTCTGCTGGTGGTGGTGGTGGTGGTGG 2308
Qy 2440 TGTCTCAATCCCT 2499
Db 2309 TGTCTCAATCCCT 2368
Qy 2500 CTGAGAAAGCAACCTAGCTCTGGGACAGATCAAAACACCCAGCTTGTATGTCAATTAAC 2559
Db 2369 CTGAGAAAGCAACCTAGCTCTGGGACAGATCAAAACACCCAGCTTGTATGTCAATTAAC 2428

Qy 2560 TCTGATGATGCGAAAAACAGTCTCTGTGACTCAACTCAAGCTTGGTAACCTTTACCAAGC 2619
Db 2429 TCTGATGATGCGAAAAACAGTCTCTGTGACTCAACTCAAGCTTGGTAACCTTTACCAAGC 2488
Qy 2620 TCCAGCATCATCTATGACCTGCTCCAGTTCGGTGCCATCACCAGCTTATCCAGTGACT 2679
Db 2489 TCCAGCATCATCTATGACCTGCTCCAGTTCGGTGCCATCACCAGCTTATCCAGTGACT 2548
Qy 2680 GAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCTCATGCTCTAA 2724
Db 2549 GAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCTCATGCTCTAA 2593
RESULT 12
LOCUS AI367723/c 374 bp mRNA linear EST 15-FEB-1999
DEFINITION qv75h01.x1 NCI CGAP Uti Homo sapiens cDNA clone IMAGE:1987441.3
similar to TR:Q90674 Q90674 LUTEINIZING HORMONE RECEPTOR ;, mRNA
sequence.
ACCESSION AI367723.1 GI:4137468
VERSION AI367723
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 374)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1807 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 307.
FEATURES
source
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1987441"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_libs="NCI-CGAP Uti"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"
ORIGIN
Query Match 13.5%; Score 369; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.6e-191;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2021 CGAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTTCTCTGCTGCTGCTGCTGCTGCA 2080
Db 374 CGAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTTCTCTGCTGCTGCTGCTGCTGCA 315
Qy 2081 CCATGGCGCGAGTTCCCTCTGCTGGTGCGACAGATGCGCCTCCCTCTCTGCTGCTGCTG 2140
Db 314 CCATGGCGCGAGTTCCCTCTGCTGGTGCGACAGATGCGCCTCCCTCTCTGCTGCTGCTG 255

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 13, 2005, 00:42:39 ; Search time 243.5 Seconds
(without alignments)
11457.115 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 907
Sequence: 1 atgacacctccggctgg.....cattgtccctgctctctaa 2724

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3224408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model
-Q=/cgn2_1/USPTO/US10751736/runat 12072005 084327 23917/app query.fasta 1.2887
-DB=UniProt_03 -QFWT=fastan -SUFFIX=oli.rup -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=pto
-USER=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	1 LGR5 HUMAN	O75473 homo sapien
2	51	5.6	363	2 Q8C8A7	Q8C8A7 mus musculus
3	51	5.6	907	1 LGR5 MOUSE	Q921P4 mus musculus
4	27	3.0	147	2 Q7IM96	Q7IM96 sus scrofa
5	18	2.0	828	1 LGR6 HUMAN	Q9Bxb8 homo sapien
6	18	2.0	915	2 Q6UY15	Q6UY15 homo sapien
7	18	2.0	923	2 Q86VU0	Q86VU0 homo sapien
8	18	2.0	928	2 Q8BYD7	Q8BYD7 homo sapien
9	15	1.7	878	2 Q8BXS9	Q8BXS9 mus musculus
10	13	1.4	927	2 Q8N537	Q8N537 homo sapien
11	13	1.4	951	1 LGR4 HUMAN	Q9Bxb1 homo sapien
12	13	1.4	951	1 LGR4 RAT	Q922H4 rattus norv
13	11	1.2	134	2 Q80T31	Q80T31 mus musculus
14	11	1.2	162	2 Q80UB8	Q80UB8 mus musculus
15	11	1.2	459	2 Q8R301	Q8R301 mus musculus
16	9	1.0	168	2 Q80Zr7	Q80Zr7 mus musculus

c

c	17	9	1.0	187	2	Q8BYF2	Q8BYF2 mus musculus
	18	9	1.0	195	2	Q8UHZ5	Q8UHZ5 agrobacteri
	19	9	1.0	209	2	Q7D1B7	Q7D1B7 agrobacteri
	20	9	1.0	213	2	Q6ANT4	Q6ANT4 desulfotale
c	21	9	1.0	218	2	Q821I3	Q821I3 streptomyc
c	22	9	1.0	232	1	RHO6 HUMAN	Q92730 homo sapien
c	23	9	1.0	232	2	Q8BLR7	Q8BLR7 mus musculus
c	24	9	1.0	232	2	Q66JC6	Q66JC6 xenopus tro
c	25	9	1.0	232	2	Q7X477	Q7X477 xenopus lae
c	26	9	1.0	232	2	Q9W761	Q9W761 xenopus lae
c	27	9	1.0	342	2	Q9XVNI	Q9XVNI caenorhabdi
c	28	9	1.0	369	2	Q93177	Q93177 caenorhabdi
	29	9	1.0	379	2	Q9DG05	Q9DG05 gallus gall
	30	9	1.0	424	2	Q9DG06	Q9DG06 gallus gall
	31	9	1.0	488	2	Q6PHA3	Q6PHA3 mus musculus
	32	9	1.0	589	2	Q9FHL8	Q9FHL8 arabidopsis
	33	9	1.0	610	2	Q7GEY4	Q7GEY4 podospora a
	34	9	1.0	638	2	Q6Z1I6	Q6Z1I6 oryza sativ
	35	9	1.0	653	2	Q8LKV9	Q8LKV9 aegilops ta
	36	9	1.0	657	2	Q6ZHK7	Q6ZHK7 oryza sativ
c	37	9	1.0	665	2	Q9F5F6	Q9F5F6 agrobacteri
	38	9	1.0	693	1	F5HR CHICK	F79763 gallus gall
	39	9	1.0	700	2	Q02714	Q02714 podospora a
	40	9	1.0	869	2	Q23253	Q23253 arabidopsis
	41	9	1.0	871	2	Q7XNV7	Q7XNV7 oryza sativ
	42	9	1.0	879	2	Q68CI3	Q68CI3 oncorhynch
	43	9	1.0	891	2	Q93YT3	Q93YT3 arabidopsis
	44	9	1.0	999	2	Q8K7X8	Q8K7X8 streptococc
	45	9	1.0	1008	2	Q8P1F7	Q8P1F7 streptococc

ALIGNMENTS

RESULT 1

LGR5_HUMAN
ID LGR5_HUMAN STANDARD; PRT; 907 AA.

AC O75473; Q9UP75;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
(Orphan G protein-coupled receptor HG38) (G protein-coupled receptor 49).

GN Name=GPR49; Synonyms=GPR67, LGR5;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98308104; PubMed=9642114; DOI=10.1006/bbrc.1998.8774;

RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;

"Identification and cloning of an orphan G protein-coupled receptor of

the glycoprotein hormone receptor subfamily.";

RL Biochem. Biophys. Res. Commun. 247:266-270(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;

RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;

"Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";

RL Mol. Endocrinol. 12:1830-1845(1998).

CC -1- FUNCTION: Orphan receptor. It may be an important receptor for

signals controlling growth and differentiation of specific

embryonic tissues (by similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal

cord, and various region of brain.

CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC -1- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

CC

Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTACCTGAACAACTAGACTCTGAATGGTGGCTCAAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTCTGATTAACTGGAACTGCAACCTGGAGAGTCTGACTTAACTTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
QY 1021 TCATCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAACTCTCAAGTCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACTATTAGAAAGATTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAGACATAATGAATCTAGAAATTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTGCTATTATTACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGACCTATCGTCCAACTCTGTGCTTTTCCCTAATACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTAGATGTTAACTCATTAAATTACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTTCAGAACTCAAGGTATTAGAAATTCCTTATGCTTTACAGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTyrAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGAGCAGCTTCATAAGAACTGCTGGAATGTTTCAAGCTCAAGATGAAGCT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCTCTGCTGACTTTTCAGGAAGACCTGGAAGCCCTTCAATCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluLeuAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCAGGCCCTTCAAAACCTGTGAAACACCTGCTTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
QY 1681 AGAATTGAGTGTGACCATGACAGTCTTGGCAGCTTACTTGTAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCAATTAACCTGTAATTTGGGGTTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCACGGAGTCTCCAGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGACAGCATGCTGGTGGTGGGGAATGGGGTGGTGGTGGTGGTGGTGGTGG 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660

QY 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAAATCATTTTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2100
Db 681 LeuLysValIleIleLeuLeuLysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuLysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCAACATGGGCTACATGCTGCTCTCATCTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCCCTACCAAGCTCTACTGCAATTTGGACAGGGAGAGCCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGTCTTTCACCAACTGCATCTCTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGCTTCTTGTCTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTCTGTGTGTAGTCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTGTGCTCAATCTCACTTTAAGGAGATCTGTGAGCCTGAGAAAGCAAACTTACCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuLysGlnThrTyrVal 840
QY 2521 TGGCAAGAGATCAAAACCAACCCAGCTTGATGTCAATTAACCTCTGTGATGTGCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGCTCACTCACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCATTTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCCGTCGTCATCACAGCTTATCCAGTGTGAGAGTGCATCTTTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTGTGCCATGCTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 2

Q8C8A7 ID Q8C8A7 PRELIMINARY; PRT; 363 AA.
AC Q8C8A7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
DE library, clone: C130018C02 product: G protein-coupled receptor 49, full
DE insert sequence.
GN Name=Gpr49;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; Pubmed=10349636; DOI=10.1016/S00076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

```
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RP [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RP [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fekuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the ENBL/GenBank/DBJ databases.
DR EMBL; AK047873; BAC33180.1; --
DR HSSP; Q9BZR6; IOZN.
DR MGD; MGI:1341817; Gpr49.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00560; LRR_1; 9.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 6.
DR Receptor.
SQ SEQUENCE 363 AA; 39879 MW; D5E2FC4449FCE2C0 CRC64;
Pred. No.: 9.1e-43 Length: 363
Alignment Scores:
DR EMBL; AF110818; AAD14684.1; --
DR HSSP; Q9BZR6; IOZN.
DR MGD; MGI:1341817; Gpr49.
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.62% Indels: 0
DB: 2 Gaps: 0
US-10-751-736-21 (1-2724) x Q8C8A7 (1-363)
Qy 607 GACTATGCTTGGAAACCTCTCAGCTTGCTAGTTCTACATCTCCATACATAGAAATC 666
Db 203 AsptYrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsnArgile 222
Qy 667 CACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGATTAAAT 726
Db 223 HisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAspLeuAsn 242
Qy 727 TACAATAACCTTGATGAATTCCTCCACTGCAATT 759
Db 243 TyrAsnAsnLeuAspGluPheProThrAlaile 253
RESULT 3
LGR5 MOUSE
ID LGR5 MOUSE STANDARD; PRT; 907 AA.
AC Q921P4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (G protein-coupled receptor 49) (Orphan G protein-coupled receptor
DE FE).
GN Name=Gpr49; Synonyms=Tex, Lgr5;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99121227; PubMed=9920770; DOI=10.1006/bbrc.1998.9882;
RA Hermey G., Methner A., Schaller H.C., Hermans-Borgmeyer I.;
RT "Identification of a novel seven-transmembrane receptor with homology
RT to glycoprotein receptors and its expression in the adult and
RT developing mouse.";
RL Biochem Biophys. Res. Commun. 254:273-279(1999).
CC -1- FUNCTION: Orphan receptor. It may be an important receptor for
CC signals controlling growth and differentiation of specific
CC embryonic tissues.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC and in the brain. In the central nervous system expression is
CC restricted to the olfactory bulb. In the adrenal gland detected
CC only in the neural-crest derived chromaffin cells of the
CC medulla, but not in the cells of the adrenal cortex. In the
CC gonads, the expression is high in Graafian follicle, but absent
CC from primary and secondary follicles.
CC -1- DEVELOPMENTAL STAGE: Expressed from embryonic day 10.5 (E10.5) in
CC the developing spinal cord and in the neuroepithelia of the
CC myel-, met-, mes-, and diencephalon. Expression is transitory and
CC the pattern changed rapidly.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -1- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF110818; AAD14684.1; --
CC HSSP; Q9BZR6; IOZN.
CC MGD; MGI:1341817; Gpr49.
```


RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.H., Yanaura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270 (2003).
DR EMBL; AY358119; AAQ8486.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Ghrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00365; LRR_SD22; 5.
DR SMART; SM00369; LRR_TYP; 14.
KW Receptor.
SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DBB555F4 CRC64;

Alignment Scores:
Pred. NO.: 1.29e-08 Length: 915
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q6UV15 (1-915)

Qy 853 ACAATACATTCTATGACAAATCCCATCAATTTGTGGAGATCTGCTTTTCAA 906
Db 233 ThrileHisPheTyraAspAsnProIleGlnPheValGlyArgSerAlaPheGln 250

RESULT 7

Q86VUO
ID AC Q86VUO PRELIMINARY; PRT; 923 AA.
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fanev J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047905; AAH47905.1; -.
DR HSPF; Q9BZR6; IPR01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Ghrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON TER 1
SQ SEQUENCE 923 AA; 100047 MW; 541D6746DAB06813 CRC64;

Alignment Scores:
Pred. NO.: 1.29e-08 Length: 923
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q86VUO (1-923)

Qy 853 ACAATACATTCTATGACAAATCCCATCAATTTGTGGAGATCTGCTTTTCAA 906
Db 241 ThrileHisPheTyraAspAsnProIleGlnPheValGlyArgSerAlaPheGln 258

RESULT 8

Q9BYD7
ID AC Q9BYD7 PRELIMINARY; PRT; 928 AA.
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE VTS20631 (Fragment).
GN Name=VTS20631;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okaze H., Hayaashi A., Kozuma S., Saito T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049405; BAB39854.1; -.
DR HSPF; P25147; IPR01.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Ghrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.


```

FT  NON TER      1      1
SQ  SEQUENCE    928 AA; 100487 MW;  4C3364ADEA89C463 CRC64;

Alignment Scores:
Pred. No.:      1.29e-08      Length:      928
Score:          18.00         Matches:    18
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.98%        Indels:    0
DB:              2           Gaps:     0

US-10-751-736-21 (1-2724) x Q9BYD7 (1-928)

QY  853 ACATACATTTCTATGACATCCATCCATCTTCTGGGAGATCGCTTTCAA 906
Db  246 ThrileHisPheTyAspAsnProIleInPheValGlyArgSerAlaPheGln 263
|||||
RESULT 9
Q8BX59 PRELIMINARY; PRT; 878 AA.
AC  Q8BX59
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE  Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DE  clone.A930009A08 product:G PROTEIN-COUPLED RECEPTOR LGR4 homolog
DE  (Fragment);
GN  Name=Gpr48;
OS  Mus musculus (Mouse);
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RX  MEDLINE=21085660; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA  Carninci P., Hayashizaki Y.;
RA  RIKEN FANTOM Consortium;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
[3]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RC  The FANTOM Consortium;
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RT  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
[4]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RC  MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA  Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA  Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT  "Normalization and subtraction of cap-trapper-selected cDNAs to
RT  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL  Genome Res. 10:1617-1630(2000).
[5]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RC  MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA  Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA  Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA  Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

```

```

RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT  "RIKEN integrated sequence analysis (RISA) system-384-format
RT  sequencing pipeline with 384 multicapillary sequencer.";
RL  Genome Res. 10:1757-1771(2000).
[6]
SEQUENCE FROM N.A.
RN  STRAIN=C57BL/6J; TISSUE=Retina;
RC  Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA  Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA  Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA  Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA  Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA  Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA  Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA  Saito R., Saitho H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA  Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA  Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA  Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK044357; BAC31882.1; -.
DR  HSSP; P25147; ID0B.
DR  MGd; MG1:1891468; Gpr48.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR  GO; GO:0004872; F:receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR  InterPro; IPR000276; GPCR_Rhodopsn.
DR  InterPro; IPR002131; Gprhnm_receptor.
DR  InterPro; IPR001611; LRR.
DR  InterPro; IPR003591; LRR typ.
DR  InterPro; IPR007087; 2nf_C2H2.
DR  Pfam; PF00001; 7tm 1; 1.
DR  Pfam; PF00560; LRR_1; 14.
DR  PRINTS; PR00373; GLYCHORMONER.
DR  PRINTS; PR00237; GPCRHHODOPSIN.
DR  PRINTS; PR00019; LEURICHRPT.
DR  SMART; SM00369; LRR TYP; 5.
DR  PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW  Receptor.
FT  NON TER      1      1
SQ  SEQUENCE    878 AA; 96877 MW;  68E64B5EDEA11B37 CRC64;

Alignment Scores:
Pred. No.:      1.66e-05      Length:      878
Score:          15.00         Matches:    15
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.65%        Indels:    0
DB:              2           Gaps:     0

US-10-751-736-21 (1-2724) x Q8BX59 (1-878)

QY  706 CTAGAGACTTTAGATTAAATTACAATTAACCTTGATCAATTCCTCC 750
Db  154 LeuGluThrLeuAspLeuAsnTyAsnAsnLeuAspGluPhePro 168
|||||
RESULT 10
Q8N537 PRELIMINARY; PRT; 927 AA.
AC  Q8N537
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  GPR48 protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
SEQUENCE FROM N.A.
RP  TISSUE=Brain;
RC  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.E., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinchwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC033039; AAB33039.1; -;
 DR HSSP; P25147; 1D0B.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003591; LRR_Nterm.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF00560; LRR_1; 14.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 5.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
 DR PROSITE; PS00028; ZINC FINGER C2H2 1; UNKNOWN 1.
 SQ SEQUENCE 927 AA; 101676 MW; C7B2F0C40E584C68 CRC64;
 Alignment Scores:
 Pred. No.: 0.00195 Length: 927
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.43% Indels: 0
 DB: 2 Gaps: 0
 US-10-751-736-21 (1-2724) x Q8NE37 (1-927)
 QY 622 AACCTCTCCAGTGGTAGTCTACATCTCCATCAACAT 660
 Db 175 AsnLeuSerLeuValValLeuHisLeuHisAsn 187
 RESULT 11
 LGR4_HUMAN
 ID_LGR4_HUMAN STANDARD; PRT; 951 AA.
 AC Q9BXL1; Q9NYD1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
 DE (G protein-coupled receptor 48).
 GN Name=GPR48; Synonyms=LGR4;
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=21294803; PubMed=11401528; DOI=10.1006/bbrc.2001.4625;
 RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
 RT "Molecular characterization of a novel glycoprotein hormone G-protein-
 coupled receptor.";
 RL Biochem. Biophys. Res. Commun. 282:757-764 (2001).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
 CC placenta, ovary, testis and adrenal. Expressed also in spinal
 CC cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
 CC and spleen.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -1- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF346711; AAK31153.1; -;
 DR EMBL; AF346709; AAK31153.1; JOINED.
 DR EMBL; AF346710; AAK31153.1; JOINED.
 DR EMBL; AF257182; AAF68989.1; -;
 DR HSSP; Q9BZR6; 1OZN.
 DR Gene; HGNC:13299; GPR48.
 DR MIN; 606666; -;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR002131; Gprhm_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR00372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_TYP.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF00560; LRR_15.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 4.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; FALSE_NEG.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 24 Potential.
 FT CHAIN 25 951 Leucine-rich repeat-containing G protein-
 FT coupled receptor 4.
 FT DOMAIN 25 544 Extracellular (Potential).
 FT TRANSMEM 545 565 1 (Potential).
 FT DOMAIN 566 575 Cytoplasmic (Potential).
 FT TRANSMEM 576 596 2 (Potential).
 FT DOMAIN 597 620 Extracellular (Potential).
 FT TRANSMEM 621 641 3 (Potential).
 FT DOMAIN 642 661 Cytoplasmic (Potential).
 FT TRANSMEM 662 682 4 (Potential).
 FT DOMAIN 683 703 Extracellular (Potential).
 FT TRANSMEM 704 724 5 (Potential).
 FT DOMAIN 725 756 Cytoplasmic (Potential).
 FT TRANSMEM 757 777 6 (Potential).
 FT DOMAIN 778 783 Extracellular (Potential).
 FT TRANSMEM 784 804 7 (Potential).
 FT DOMAIN 805 951 Cytoplasmic (Potential).
 FT REPEAT 55 79 LRR 1.
 FT REPEAT 81 103 LRR 2.
 FT REPEAT 104 127 LRR 3.

```

FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 201 223 LRR 7.
FT REPEAT 224 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 368 387 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 By similarity.
FT CARBOHYD 68 68 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc... ) (Potential).
FT CONFLICT 292 292 P -> S (in Ref. 1; AAF68989).
FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Alignment Scores:
Pred. No.: 0.00195 Length: 951
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-21 (1-2724) x LGR4_HUMAN (1-951)
Qy 622 AACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 199 AsnLeuSerLeuValValLeuHisLeuHisAsnAen 211

RESULT 12
LGR4_RAT STANDARD; PRT; 951 AA.
AC Q92ZHA;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN Name=Gpr48; Synonyms=Lgr4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
RA Hsu S.Y., Liang S.-G., Heueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF061443; AAC77910.1; --

```

```

DR HSP: Q9BZR6; LOZN.
DR RGD; 628615; Gp48.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR00372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 15.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHROMONER.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PRINTS; PR00019; LEURICHREPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECPT_F1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
KW SIGNAL 1 24 Potential.
FT CHAIN 25 951 Leucine-rich repeat-containing G protein-
FT FT Extracellular (Potential).
FT DOMAIN 25 544 Extracellular (Potential).
FT TRANSMEM 545 565 1 (Potential).
FT DOMAIN 566 575 Cytoplasmic (Potential).
FT TRANSMEM 576 596 2 (Potential).
FT DOMAIN 597 619 Extracellular (Potential).
FT TRANSMEM 620 640 3 (Potential).
FT DOMAIN 641 661 Cytoplasmic (Potential).
FT TRANSMEM 662 682 4 (Potential).
FT DOMAIN 683 703 Extracellular (Potential).
FT TRANSMEM 704 724 5 (Potential).
FT DOMAIN 725 756 Cytoplasmic (Potential).
FT TRANSMEM 757 777 6 (Potential).
FT DOMAIN 778 783 Extracellular (Potential).
FT TRANSMEM 784 804 7 (Potential).
FT DOMAIN 805 951 Cytoplasmic (Potential).
FT REPEAT 55 79 LRR 1.
FT REPEAT 81 103 LRR 2.
FT REPEAT 104 127 LRR 3.
FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 200 223 LRR 7.
FT REPEAT 225 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 368 387 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 By similarity.
FT CARBOHYD 68 68 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 188 188 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Alignment Scores:
Pred. No.: 0.00195 Length: 951
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-21 (1-2724) x LGR4_RAT (1-951)
Qy 622 AACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660

```

Db 199 AsnLeuSerLeuValValLeuHisLeuHisAsnAen 211

RESULT 13

Q80T31 PRELIMINARY; PRT; 134 AA.

AC Q80T31; 0-283 Length: 162

DT 01-JUN-2003 (TrEMBLrel. 24, Created) Matches: 11

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update) Conservative: 0

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update) Mismatches: 0

DE G protein-coupled receptor GPR48 (Fragment). Indels: 0

OS Mus musculus (Mouse). Gaps: 0

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;

RA Vassilatis D.K., Hohmann J.G., Zeng H., Li P., Ranchalis J.E.,

RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,

RA Bergmann J.E., Gaitanaris G.A.;

RT "The G protein-coupled receptor repertoire of human and mouse.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).

DR EMBL; AY255619; AA085131.1; --

DR HSP; P25146; 10FT.

DR GO; GO:0004872; F:receptor activity; IEA.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003591; LRR typ.

DR Pfam; PF00560; LRR_1; 4.

DR PRINTS; PR00019; LEURICRPT.

DR SMART; SM00369; LRR_TYP; 2.

KW Receptor.

FT NON TER 1 1

FT NON TER 134 134

SQ SEQUENCE 134 AA; 14851 MW; 4D355900D69C67E CRC64;

Alignment Scores:

Pred. No.: 0.289 Length: 134

Score: 11.00 Matches: 11

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.21% Indels: 0

DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q80T31 (1-134)

Qy 418 CTTCAATCCGCTGCTGATGCTAACACATC 450

Db 69 LeuGlnSerLeuArgLeuAspAlaAsnHisile 79

RESULT 14

Q80UB8 PRELIMINARY; PRT; 162 AA.

AC Q80UB8; 0-289 Length: 162

DT 01-JUN-2003 (TrEMBLrel. 24, Created) Matches: 11

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update) Conservative: 0

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update) Mismatches: 0

DE Leucine-rich repeat-containing G protein-coupled receptor 6 (Fragment). Indels: 0

OS Mus musculus (Mouse). Gaps: 0

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;

RA Vassilatis D.K., Hohmann J.G., Zeng H., Li P., Ranchalis J.E.,

RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,

RA Bergmann J.E., Gaitanaris G.A.;

RT "The G protein-coupled receptor repertoire of human and mouse.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).

DR EMBL; AY255562; AA085074.1; --

DR GO; GO:0004872; F:receptor activity; IEA.

KW Receptor.

FT NON TER 1 1

FT NON TER 162 162

SQ SEQUENCE 162 AA; 17608 MW; C593128C551FB824 CRC64;

Alignment Scores:

Pred. No.: 0.283 Length: 162

Score: 11.00 Matches: 11

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.21% Indels: 0

DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q80UB8 (1-162)

Qy 2428 CCACCTTCGATGCTCAATCCCTTCTCTAC 2460

Db 70 ProLeuProAlaCysLeuAenProLeuLeuTyr 80

RESULT 15

Q8R301 PRELIMINARY; PRT; 459 AA.

AC Q8R301; 0-283 Length: 162

DT 01-JUN-2002 (TrEMBLrel. 21, Created) Matches: 11

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update) Conservative: 0

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update) Mismatches: 0

DE Lgr6 protein (Fragment). Indels: 0

GN Name=Lgr6;

OS Mus musculus (Mouse). Gaps: 0

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Czech II; TISSUE=Mammary tumor; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smallos D.B., Schnerch A., Schein J.E.,

RA Jones S.J., Marta M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=Czech II; TISSUE=Mammary tumor;

RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC026896; AAH26896.1; --

DR MGI; MGI:2441805; Lgr6.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.

DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR InterPro; IPR002131; Gphrmn_receptor.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00373; GLYCHORMONER.

DR PRINTS; PR00237; GPCRHHODPSN.

FT NON TER 1 1

FT NON TER 459 AA; 47889 MW; F0100BF073E81762 CRC64;

Alignment Scores:
Pred. No.: 0.251 Length: 459
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q8R301 (1-459)

Qy 2428 CCACTTCCTGCATGCTCAATCCCTTCTCTAC 2460
Db 309 ProLeuProLaCysLeuAsnProLeuTyr 319

Search completed: July 13, 2005, 03:56:52
Job time : 284.5 secs

19	799.4	29.3	3381	6	AX527920	Se
----	-------	------	------	---	----------	----

181	Db	CTTCCAACTCAGCGTCTTCCACCTCTTACCTCTAGACCTCAGTATGAACAACATCAGTCAG	240
241	QY	CTGCTCCCGAATCCCTGCGCAGTCTCGCTCTCTGGAGGAGTTACGCTTTCGGGGAAAC	300
241	Db	CTGCTCCCGAATCCCTGCGCAGTCTCGCTCTCTGGAGGAGTTACGCTTTCGGGGAAAC	300
301	QY	GCTCTGACATACATTTCCCAAGGAGCATTCACTGGCCCTTTACAGCTTTAAAGTTCTTTATG	360
301	Db	GCTCTGACATACATTTCCCAAGGAGCATTCACTGGCCCTTTACAGCTTTAAAGTTCTTTATG	360
361	QY	CTGCAGAAATATCAGTCTAAGACA	420
361	Db	CTGCAGAAATATCAGTCTAAGACA	420
421	QY	CAATCCCTCGCTCTGGATGTAA	480
421	Db	CAATCCCTCGCTCTGGATGTAA	480
481	QY	CTGCATTTCCCTGAGGCACCTCTGGCTGGAATGCAATCGTTTAA	540
481	Db	CTGCATTTCCCTGAGGCACCTCTGGCTGGAATGCAATCGTTTAA	540
541	QY	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAA	600
541	Db	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAA	600
601	QY	ATACGAGATATGCTTTGGGAAACCTCTCAGCTGTGTAGTTCTACATCTCCATAA	660
601	Db	ATACGAGATATGCTTTGGGAAACCTCTCAGCTGTGTAGTTCTACATCTCCATAA	660
661	QY	AGAAATCCACTCCCTGGGAAAGAAATGCTTTCATGGGCTCCACAGCTAGAGACTTTAGAT	720
661	Db	AGAAATCCACTCCCTGGGAAAGAAATGCTTTCATGGGCTCCACAGCTAGAGACTTTAGAT	720
721	QY	TTAAATTAACAATAACCTTTGTAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTTAA	780
721	Db	TTAAATTAACAATAACCTTTGTAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTTAA	780
781	QY	GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAGCATTTGTAGGCAAC	840
781	Db	GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAGCATTTGTAGGCAAC	840
841	QY	CCTTCTCTTATTAACAATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGACTCTGCT	900
841	Db	CCTTCTCTTATTAACAATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGACTCTGCT	900
901	QY	TTTCAACAATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCTCA	960
901	Db	TTTCAACAATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCTCA	960
961	QY	TTTCTGTATTTAACTGGAACCTGGAACCTGGAAGTCTGACTTTAACTGGAGCAGATC	1020
961	Db	TTTCTGTATTTAACTGGAACCTGGAACCTGGAAGTCTGACTTTAACTGGAGCAGATC	1020
1021	QY	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGCT	1080
1021	Db	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGCT	1080
1081	QY	TACAACCTTATTAGAAGATTTACCCAGTTTTCCTGCTGCCCCAAAGCTTCAGAAAATTTGAC	1140
1081	Db	TACAACCTTATTAGAAGATTTACCCAGTTTTCCTGCTGCCCCAAAGCTTCAGAAAATTTGAC	1140
1141	QY	CTAAGACATAATGAAATCTAGAAATTAAGTTGACATTTCCAGCAGTGTGCTTAGGCCTC	1200
1141	Db	CTAAGACATAATGAAATCTAGAAATTAAGTTGACATTTCCAGCAGTGTGCTTAGGCCTC	1200
1201	QY	CGATCGCTGAAATTTGGCTTGGAACAAATTCGCTATTTATCCCCCAATGCAATTTCCACT	1260
1201	Db	CGATCGCTGAAATTTGGCTTGGAACAAATTCGCTATTTATCCCCCAATGCAATTTCCACT	1260
1261	QY	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTGCTCTTTTCTTATAACT	1320

13261	DB	TTGCCATCCCTTAATAAAGCTGGGACCTTATCGTCCAACTCCTCTGCTGCTTTTCTCTTAATAACT	13262
1321	QY	GGGTTACATGGTTTAACTCACTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGCATA	1380
1321	DB	GGGTTACATGGTTTAACTCACTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGCATA	1380
1381	QY	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATATAGAAATGCTTATGCTTATACCAAGTGTCTGT	1440
1381	DB	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATATAGAAATGCTTATGCTTATACCAAGTGTCTGT	1440
1441	QY	GCATTTGGAGTGTGTGAGATGCTTATAGATTTCTTAATCAAATGGAATTAAGGTGACAC	1500
1441	DB	GCATTTGGAGTGTGTGAGATGCTTATAGATTTCTTAATCAAATGGAATTAAGGTGACAC	1500
1501	QY	AGCAGTATGACGACCTTCTATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
1501	DB	AGCAGTATGACGACCTTCTATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
1561	QY	GACCTTGAAGAATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
1561	DB	GACCTTGAAGAATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
1621	QY	TGTTCACTTCCCAGGCCCCCTTCAAAACCCTGTGAACACCTGCTTGAATGCTGGCTGGCTGATC	1680
1621	DB	TGTTCACTTCCCAGGCCCCCTTCAAAACCCTGTGAACACCTGCTTGAATGCTGGCTGGCTGATC	1680
1681	QY	AGAAATCGAGTGTGACCATAGCAGTGTCTGGCACTTACTTTGTAATGCTTTGGTGGACTTCA	1740
1681	DB	AGAAATCGAGTGTGACCATAGCAGTGTCTGGCACTTACTTTGTAATGCTTTGGTGGACTTCA	1740
1741	QY	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCACTTAAACTGTTAAATGGGGTCATCGCA	1800
1741	DB	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCACTTAAACTGTTAAATGGGGTCATCGCA	1800
1801	QY	GCAGTGAACATGCTCAGCGGAGTCTCAAGTCCGCTGCTGGCTGGTGGATGGGTTCACT	1860
1801	DB	GCAGTGAACATGCTCAGCGGAGTCTCAAGTCCGCTGCTGGCTGGTGGATGGGTTCACT	1860
1861	QY	TTTGGCAGCTTTGCACACATGCTGCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCAATT	1920
1861	DB	TTTGGCAGCTTTGCACACATGCTGCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCAATT	1920
1921	QY	GGTTTTTTTGCCATTTTGTGCTTCCAGAAATCACTGTCTTCTGCTTACTCTGGCAGGCCCTG	1980
1921	DB	GGTTTTTTTGCCATTTTGTGCTTCCAGAAATCACTGTCTTCTGCTTACTCTGGCAGGCCCTG	1980
1981	QY	GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040
1981	DB	GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040
2041	QY	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCTTGAACATGGCCGAGTTCCCTCG	2100
2041	DB	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCTTGAACATGGCCGAGTTCCCTCG	2100
2101	QY	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTATGCTTTTGGGGAGCCCC	2160
2101	DB	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTATGCTTTTGGGGAGCCCC	2160
2161	QY	AGCACCATGGGCTACATGGTTCATCTGCTTCAATTTCCCTTTGCTTCTCCTCATGATG	2220
2161	DB	AGCACCATGGGCTACATGGTTCATCTGCTTCAATTTCCCTTTGCTTCTCCTCATGATG	2220
2221	QY	ACCAATGCCCTACACCAAGCTCTACTGCAATTTTGGAACAGGGAGACCTTGGAGAAATATTGG	2280
2221	DB	ACCAATGCCCTACACCAAGCTCTACTGCAATTTTGGAACAGGGAGACCTTGGAGAAATATTGG	2280
2281	QY	GACTGCTCTATGTGTAAACACATTTGCGCTCTTGTCTTCCACCAATGCATCTCTAAACTGC	2340
2281	DB	GACTGCTCTATGTGTAAACACATTTGCGCTCTTGTCTTCCACCAATGCATCTCTAAACTGC	2340
2341	QY	CCGTGGGCTTCTTGTCTCTCTTTTAATAAACCTTACATTTATCAAGTCTCTGAAGTA	2400
2341	DB	CCGTGGGCTTCTTGTCTCTCTTTTAATAAACCTTACATTTATCAAGTCTCTGAAGTA	2400

1321	Db		GGGTTACATGGTTTTAACTCACTTAAAAATAACAGGAATCATGCTTTACAGAGCTTGATA	1380
1381	Qy	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1440	
1381	Db	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1440	
1441	Qy	GCATTTGGAGTGTGTGAGATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1500	
1441	Db	GCATTTGGAGTGTGTGAGATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1500	
1501	Qy	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560	
1501	Db	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560	
1561	Qy	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG	1620	
1561	Db	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG	1620	
1621	Qy	TGTTTCACTTCCCAGGCCCCCTTCAAACCCCTGTGAACAACCTGCTTGATGCTGGCTGATC	1680	
1621	Db	TGTTTCACTTCCCAGGCCCCCTTCAAACCCCTGTGAACAACCTGCTTGATGCTGGCTGATC	1680	
1681	Qy	AGAAATGGAGTGTGGACAATAGCAGTTCCTGGCACTTACTTGTAACTTTGGTGTACTTCA	1740	
1681	Db	AGAAATGGAGTGTGGACAATAGCAGTTCCTGGCACTTACTTGTAACTTTGGTGTACTTCA	1740	
1741	Qy	ACAGTTTTTCAGATCCCCCTCTGTACATTTTCCCCCAATTAACATGTTAATGGGTGATCGCA	1800	
1741	Db	ACAGTTTTTCAGATCCCCCTCTGTACATTTTCCCCCAATTAACATGTTAATGGGTGATCGCA	1800	
1801	Qy	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGGTTCACCT	1860	
1801	Db	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGGTTCACCT	1860	
1861	Qy	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTGGTTCGCCATGTCATT	1920	
1861	Db	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTGGTTCGCCATGTCATT	1920	
1921	Qy	GGTTTTTTTGTCCATTTTTTGTCTTCCAGATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980	
1921	Db	GGTTTTTTTGTCCATTTTTTGTCTTCCAGATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980	
1981	Qy	GAGCGTGGTTCTCTGTGAAATATCTCGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040	
1981	Db	GAGCGTGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040	
2041	Qy	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGCTGSCCTTGAACATGGCGCAGTTTCCCCTG	2100	
2041	Db	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGCTGSCCTTGAACATGGCGCAGTTTCCCCTG	2100	
2101	Qy	CTGGGTGGCAGCATGATGAGGCGCTCCCTCTCTGCCCTCTGCTTGGCCCTTTGGGAGCCC	2160	
2101	Db	CTGGGTGGCAGCATGATGAGGCGCTCCCTCTCTGCCCTCTGCTTGGCCCTTTGGGAGCCC	2160	
2161	Qy	AGCACCATGGCTACATGGTCGCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG	2220	
2161	Db	AGCACCATGGCTACATGGTCGCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG	2220	
2221	Qy	ACCATTTGCTACACCAAGCTCTACTGCAATTTGGAACGAGGAGACCTGGAGAAATATTGG	2280	
2221	Db	ACCATTTGCTACACCAAGCTCTACTGCAATTTGGAACGAGGAGACCTGGAGAAATATTGG	2280	
2281	Qy	GACTGCTCTATGGTAAAAACAATTTGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC	2340	
2281	Db	GACTGCTCTATGGTAAAAACAATTTGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC	2340	
2341	Qy	CCTGTGGCTTTCTGTCTCTCTCTCTTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA	2400	
2341	Db	CCTGTGGCTTTCTGTCTCTCTCTCTTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA	2400	
2401	Qy	ATTAAGTTTATCTTCTTGGTGGTAGTCCCACCTTCTCGCATGTCTCAATCCCCTTCTCTAC	2460	

[illegible]

361	Db		CTGCAGAAATAATCAGCTTAAGACACGTTACCCACAGAAAGCTCTGCGAGAATTTGCGGAAGCGTT	420
421	Qy	CAATCCCTCGCTCTGGATGCTTAACACATCAGCTATGTTGCCCCCAAGCTGTTTCAGTGGC	480	
421	Db	CAATCCCTCGCTCTGGATGCTTAACACATCAGCTATGTTGCCCCCAAGCTGTTTCAGTGGC	480	
481	Qy	CTGCATTTCCCTGAGGCACCTGTGGTGGATGACAAATGCGTTTAAACAGAAATCCCGTCAG	540	
481	Db	CTGCATTTCCCTGAGGCACCTGTGGTGGATGACAAATGCGTTTAAACAGAAATCCCGTCAG	540	
541	Qy	GCITTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGGAACAAAATACACAC	600	
541	Db	GCITTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGGAACAAAATACACAC	600	
601	Qy	ATACACAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTAGTTCCTACATCTCCATAACAAT	660	
601	Db	ATACACAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTAGTTCCTACATCTCCATAACAAT	660	
661	Qy	AGAAATCCATCTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720	
661	Db	AGAAATCCATCTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720	
721	Qy	TTAAATTTACAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTTAAA	780	
721	Db	TTAAATTTACAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTTAAA	780	
781	Qy	GAACTAGGATTTTATAGCAACAAATATCAGGTCGATACCTCGAAGACATTTGTAGGCAAC	840	
781	Db	GAACTAGGATTTTATAGCAACAAATATCAGGTCGATACCTCGAAGACATTTGTAGGCAAC	840	
841	Qy	CTTTCTCTTATTTACAATACATTTTCTATGACAAATCCCATCCAATTTGTTGGAGACTGCT	900	
841	Db	CTTTCTCTTATTTACAATACATTTTCTATGACAAATCCCATCCAATTTGTTGGAGACTGCT	900	
901	Qy	TTTCAAACATTTTACTGAACTAAGAACACTGACTCTGAAATGGTGCTCACAAAATACGTAA	960	
901	Db	TTTCAAACATTTTACTGAACTAAGAACACTGACTCTGAAATGGTGCTCACAAAATACGTAA	960	
961	Qy	TTTCTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGACAGATC	1020	
961	Db	TTTCTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGACAGATC	1020	
1021	Qy	TCATCTCTTCTCTCAAACCGTCTGCAATCAGTTTACTTAATCTCCAAGTGTAGATCTGTCT	1080	
1021	Db	TCATCTCTTCTCTCAAACCGTCTGCAATCAGTTTACTTAATCTCCAAGTGTAGATCTGTCT	1080	
1081	Qy	TACAACCTTATAGAAATTTTACCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140	
1081	Db	TACAACCTTATAGAAATTTTACCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140	
1141	Qy	CTAAGCATATGAATTTAAGTAAAGTTGACACTTTTCCAGCAGTTGCTTAGGCCT	1200	
1141	Db	CTAAGCATATGAATTTAAGTAAAGTTGACACTTTTCCAGCAGTTGCTTAGGCCT	1200	
1201	Qy	CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCACCCCAATGCATTTTCCACT	1260	
1201	Db	CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCACCCCAATGCATTTTCCACT	1260	
1261	Qy	TTGCCATCCCTTAATAAGCTGGACCTATPGTCCAACTCTCTGTGCTTTTCTTAATACT	1320	
1261	Db	TTGCCATCCCTTAATAAGCTGGACCTATPGTCCAACTCTCTGTGCTTTTCTTAATACT	1320	
1321	Qy	GGGTTACATGGTTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA	1380	
1321	Db	GGGTTACATGGTTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA	1380	
1381	Qy	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAAGTGTGT	1440	
1381	Db	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAAGTGTGT	1440	
1441	Qy	GCATTTGGAGTGTGTGAATGTCCTTAAGATTTTCTAATCAATGGAATAAAGGTGACAAC	1500	

Qy	1561	GACCTTGAAGATTTCCTGCTTGACTTTTGAGGAGAGACCTGAAAGCCCTTCAATTCAAGTGCAG	1620
Db	1561	GACCTTGAAGATTTCCTGCTTGACTTTTGAGGAGAGACCTGAAAGCCCTTCAATTCAAGTGCAG	1620
Qy	1621	TGTTCACTTCCCCAGAGCCCTTCAAAACCTGTGGAACACTGCTGTGATGCTGGCTGGCTGATC	1680
Db	1621	TGTTCACTTCCCCAGAGCCCTTCAAAACCTGTGGAACACTGCTGTGATGCTGGCTGGCTGATC	1680
Qy	1681	AGAAATGGAGTGTGGACATAGCAGTTCTGGCAGCTTACTTTGTAATGTTCTTTGGTGACTTCA	1740
Db	1681	AGAAATGGAGTGTGGACATAGCAGTTCTGGCAGCTTACTTTGTAATGTTCTTTGGTGACTTCA	1740
Qy	1741	ACAGTTTTCAGATCCCCTCTGTACATTTTCCCCCATTAATACTGTTAAATGCGGGTCATGCGCA	1800
Db	1741	ACAGTTTTCAGATCCCCTCTGTACATTTTCCCCCATTAATACTGTTAAATGCGGGTCATGCGCA	1800
Qy	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGCTGGCTGGTGTGAGATGCGTTCACT	1860
Db	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGCTGGCTGGTGTGAGATGCGTTCACT	1860
Qy	1861	TTTGGCAGCTTTGACGACATGGTGCCTGTGGGAGAAATGGGTTGGTGGCCATGTCATT	1920
Db	1861	TTTGGCAGCTTTGACGACATGGTGCCTGTGGGAGAAATGGGTTGGTGGCCATGTCATT	1920
Qy	1921	GGTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	1980
Db	1921	GGTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	1980
Qy	1981	GAGCGTGGGTTCTCTGTGAAATAATTCTGCAAAATTTGAAAAGCTCAATTTTCTTAGC	2040
Db	1981	GAGCGTGGGTTCTCTGTGAAATAATTCTGCAAAATTTGAAAAGCTCAATTTTCTTAGC	2040
Qy	2041	CTGAAAGTAATCATTTTGTCTCTGCGCCTGTGGCCTTTGACCATGGCCGAGTTCCCTCG	2100
Db	2041	CTGAAAGTAATCATTTTGTCTCTGCGCCTGTGGCCTTTGACCATGGCCGAGTTCCCTCG	2100
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCGCTGCTTTGCGTTTGGGGAGCCCC	2160
Db	2101	CTGGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCGCTGCTTTGCGTTTGGGGAGCCCC	2160
Qy	2161	AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTTCCCTTGTCTTCATGATG	2220
Db	2161	AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTTCCCTTGTCTTCATGATG	2220
Qy	2221	ACCAATTGCCCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGG	2280
Db	2221	ACCAATTGCCCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGG	2280
Qy	2281	GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTACCAACTGCATCTCTAAACTGC	2340
Db	2281	GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTACCAACTGCATCTCTAAACTGC	2340
Qy	2341	CCTGTGGCTTCTTGTCCTTCTCCTCTTTAATAAACCTTACAATTATCACTGCTCGAAGTA	2400
Db	2341	CCTGTGGCTTCTTGTCCTTCTCCTCTTTAATAAACCTTACAATTATCACTGCTCGAAGTA	2400
Qy	2401	ATTAAAGTTTATCCTTCTGGTGGTAGTCCCACTTCTCTGCATGCTCAATCCCTTCTCTAC	2460
Db	2401	ATTAAAGTTTATCCTTCTGGTGGTAGTCCCACTTCTCTGCATGCTCAATCCCTTCTCTAC	2460
Qy	2461	ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAAGGAGTCTGATGATGCGAAGAACTAGCTC	2520
Db	2461	ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAAGGAGTCTGATGATGCGAAGAACTAGCTC	2520
Qy	2521	TGGAACAAGATCAAAACACCCCAAGCTTGTGATCAATTAACCTCTGATGATGTCGAAAAACAG	2580
Db	2521	TGGAACAAGATCAAAACACCCCAAGCTTGTGATCAATTAACCTCTGATGATGTCGAAAAACAG	2580
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTTG	2640
Db	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTTG	2640

Qy	2641	CTCTCCAGTTC	CGTGC	CATCAC	CAGCTTAT	CCAGTGAGCTG	GCATCTTTCTCT	2700
Db	2641	CTCTCCAGTTC	CGTGC	CATCAC	CAGCTTAT	CCAGTGAGCTG	GCATCTTTCTCT	2700
Qy	2701	GTGGCATTT	GTCCCAT	GTCTCTAA	2724			
Db	2701	GTGGCATTT	GTCCCAT	GTCTCTAA	2724			
RESULT 5								
AX658241								
LOCUS	AX658241			2880 bp	DNA	linear	PAT 22-MAR-2003	
DEFINITION	Sequence 157 from Patent WO03000928.							
ACCESSION	AX658241							
VERSION	AX658241.1			GI:29160776				
KEYWORDS								
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
AUTHORS	1 Poulsen, H.S., Pedersen, N., Mortensen, S., Sorensen, S.B., Petersen, M.W. and Elsenner, H.I.							
TITLE	Methods for identification of cancer cell surface molecules and cancer specific promoters, and therapeutic uses thereof							
JOURNAL	Patent: WO 03000928-A 157 03-JAN-2003;							
	Odin Medical A/S (DK)							
FEATURES	Location/Qualifiers							
source	1..2880							
	/organism="Homo sapiens"							
	/mol_type="unassigned DNA"							
	/db_xref="taxon:9606"							
ORIGIN								
Query Match	100.0%; Score 2724; DB 6; Length 2880;							
Best Local Similarity	100.0%; Pred. No. 0;							
Matches 2724; Conservative	0; Mismatches 0; Indels 0; Gaps 0;							
Qy	1	ATG	GAC	ACC	TCC	CGG	GCT	CGT
Db	49	ATG	GAC	ACC	TCC	CGG	GCT	CGT
Qy	61	G	G	G	G	G	G	G
Db	109	G	G	G	G	G	G	G
Qy	121	G	A	C	C	G	G	A
Db	169	G	A	C	C	G	G	A
Qy	181	C	T	T	C	A	A	C
Db	229	C	T	T	C	A	A	C
Qy	241	C	T	G	T	C	C	G
Db	289	C	T	G	T	C	C	G
Qy	301	G	C	T	C	A	T	A
Db	349	G	C	T	C	A	T	A
Qy	361	C	T	C	A	G	A	T
Db	409	C	T	C	A	G	A	T
Qy	421	C	A	A	T	C	C	T
Db	469	C	A	A	T	C	C	T
Qy	481	C	T	C	A	T	T	C
Db	529	C	T	C	A	T	T	C

Qy	541	GCTTTTGAAGTTTATCGGCATTTCGAAGCCATGACCTTGGCCCTGGAACAAAAATACACCAC	600
Db	589	GCTTTTGAAGTTTATCGGCATTTCGAAGCCATGACCTTGGCCCTGGAACAAAAATACACCAC	648
Qy	601	ATACGAGCATATGCTTTGGNAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	660
Db	649	ATACGAGCATATGCTTTGGNAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	708
Qy	661	AGAACTCACCTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Db	709	AGAACTCACCTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	768
Qy	721	TTAAATTAACAATAACCTTTGATGAATTCGCCACTGCAATTAGGACACATCTCCAACTTAAA	780
Db	769	TTAAATTAACAATAACCTTTGATGAATTCGCCACTGCAATTAGGACACATCTCCAACTTAAA	828
Qy	781	GAACCTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db	829	GAACCTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	888
Qy	841	CTTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT	900
Db	889	CTTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGGAGATCTGCT	948
Qy	901	TTTCCAAACATTTACCTGGAACCTAAGAAACACTGACTCTGAAATGGTGCTCCACAAATAACTGAA	960
Db	949	TTTCCAAACATTTACCTGGAACCTAAGAAACACTGACTCTGAAATGGTGCTCCACAAATAACTGAA	1008
Qy	961	TTTCTCTGATTTAACTGGAACCTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db	1009	TTTCTCTGATTTAACTGGAACCTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1068
Qy	1021	TCATCTCTTCTGAAAACGGTCTGCAATCAGTTTACCTTAATCTCCAGTGCTTAGATCTGCT	1080
Db	1069	TCATCTCTTCTGAAAACGGTCTGCAATCAGTTTACCTTAATCTCCAGTGCTTAGATCTGCT	1128
Qy	1081	TACAACCTATTAGAAAGATTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140
Db	1129	TACAACCTATTAGAAAGATTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1188
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1200
Db	1189	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1248
Qy	1201	CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTCCACT	1260
Db	1249	CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTCCACT	1308
Qy	1261	TTGCGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCTGTCGTCTTTTCCCTATAAAT	1320
Db	1309	TTGCGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCTGTCGTCTTTTCCCTATAAAT	1368
Qy	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1369	GGGTTACATGGTTTAACTCACTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1428
Qy	1381	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCGAGTCTGT	1440
Db	1429	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCGAGTCTGT	1488
Qy	1441	GCATTTCCGAGTGTGTGAGATGCCTATATGAAATTTCTAATCAATGGAATTAAGGTGACAAC	1500
Db	1489	GCATTTCCGAGTGTGTGAGATGCCTATATGAAATTTCTAATCAATGGAATTAAGGTGACAAC	1548
Qy	1501	AGCAGTATGGACGACCTTTCATAAGAAAAGATGCTGGAATGTTTTCAAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGACGACCTTTCATAAGAAAAGATGCTGGAATGTTTTCAAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAAGCCCTTCAATCAGTCCAG	1620
Db	1609	GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAAGCCCTTCAATCAGTCCAG	1668

Qy	1621	TGTTCACTTCCGCCAGGCCCTTCAAACCGCTGTGAAACACTGCTTGATGGCTGGCTGATC	1688
Db	1669	TGTTCACTTCCGCCAGGCCCTTCAAACCGCTGTGAAACACTGCTTGATGGCTGGCTGATC	1728
Qy	1681	AGAAATGGAGTGGGACCAATAGCAGATTCTGGCACCTTACTTGTAATGCTTTGGTGACTTCA	1740
Db	1729	AGAAATGGAGTGGGACCAATAGCAGATTCTGGCACCTTACTTGTAATGCTTTGGTGACTTCA	1788
Qy	1741	ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCATTAATACTGTTAATGGGGTCATCGCA	1800
Db	1789	ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCATTAATACTGTTAATGGGGTCATCGCA	1848
Qy	1801	GCAGTGAACATGCTCACGGAGTCTCAGTGGCGTCTGGCTGGTGTGTGATGCTTTCACT	1860
Db	1849	GCAGTGAACATGCTCACGGAGTCTCAGTGGCGTCTGGCTGGTGTGTGATGCTTTCACT	1908
Qy	1861	TTTTGGCAGCTTTCGACGACATGGTGTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGCAATT	1920
Db	1909	TTTTGGCAGCTTTCGACGACATGGTGTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGCAATT	1968
Qy	1921	GGTTTTTTTGTCCATTTTTTGCTTCAGAAATCAATCTGTTTTTCTGTCTTACTCTGGCAGGCCCTG	1980
Db	1969	GGTTTTTTTGTCCATTTTTTGCTTCAGAAATCAATCTGTTTTTCTGTCTTACTCTGGCAGGCCCTG	2028
Qy	1981	GAGCGTGGGTTCCTGTGAAATAATCTGCAAAATTTGGAACGAAAGCTCAATTTTCTAGC	2040
Db	2029	GAGCGTGGGTTCCTGTGAAATAATCTGCAAAATTTGGAACGAAAGCTCAATTTTCTAGC	2088
Qy	2041	CTGAAAGTAATCAATTTTGTCTCTGTGCGCTGCTGGCCCTTGACCAATGGCCGACGTTCCCTCG	2100
Db	2089	CTGAAAGTAATCAATTTTGTCTCTGTGCGCTGCTGGCCCTTGACCAATGGCCGACGTTCCCTCG	2148
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGCTTTTGGGGAGCCCT	2160
Db	2149	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGCTTTTGGGGAGCCCT	2208
Qy	2161	AGCACATGGGCTACATGCTGCTCATCTTGTCTCAAATTCCTTTGCTTCTCTCATGATG	2220
Db	2209	AGCACATGGGCTACATGCTGCTCATCTTGTCTCAAATTCCTTTGCTTCTCTCATGATG	2268
Qy	2221	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTCGAGAAATATTTGG	2280
Db	2269	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTCGAGAAATATTTGG	2328
Qy	2281	GACTGCTCTATGTGTAAACAACATTTGCCCTGTGTCTTCAACCAATGCACTCTTAACACTGC	2340
Db	2329	GACTGCTCTATGTGTAAACAACATTTGCCCTGTGTCTTCAACCAATGCACTCTTAACACTGC	2388
Qy	2341	CCTGTGGCTTTCTGTCTTCTCTTTTAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2400
Db	2389	CCTGTGGCTTTCTGTCTTCTCTTTTAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2448
Qy	2401	ATTTAAGTTTATCTCTTGTGGTGTAGTCCCACTTCTCGCATGCTCTCAATCCCTTCTCTAC	2460
Db	2449	ATTTAAGTTTATCTCTTGTGGTGTAGTCCCACTTCTCGCATGCTCTCAATCCCTTCTCTAC	2508
Qy	2461	ATCTTGTTCATCTCCTCACTTTTAAGGAGGATCTGTGTGAGCTCTGAGAAAGCAAACTACGTC	2520
Db	2509	ATCTTGTTCATCTCCTCACTTTTAAGGAGGATCTGTGTGAGCTCTGAGAAAGCAAACTACGTC	2568
Qy	2521	TGGACAAGATCAAAACAACCAAGCTTTGATGTCAATTTAACTCTGATGATGTGCAAAAACAG	2580
Db	2569	TGGACAAGATCAAAACAACCAAGCTTTGATGTCAATTTAACTCTGATGATGTGCAAAAACAG	2628
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACCAGCTCCAGCATCACTTATGACCTTG	2640
Db	2629	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACCAGCTCCAGCATCACTTATGACCTTG	2688
Qy	2641	CCTTCCAGTTTCCGTGCCATCACCAGCTTATCCAGTGACTTGAGAGCTGCCATCTTTTCTCT	2700
Db	2689	CCTTCCAGTTTCCGTGCCATCACCAGCTTATCCAGTGACTTGAGAGCTGCCATCTTTTCTCT	2748
Qy	2701	GTGGCAATTTGTGCCATGTCTCTTAA	2724

```
Db      2749 GTGGCATTGTCCTCCATGCTCTAA 2772
|||||
RESULT 6
AF062006
LOCUS   Homo sapiens orphan G protein-coupled receptor HG38 mRNA, complete
DEFINITION
cds.
ACCESSION AF062006
VERSION   AF062006.1 GI:3366801
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          McDonald, T., Wang, R., Bailey, W., Xie, G., Chen, F., Caskey, C.T. and
          Liu, Q.
          Identification and cloning of an orphan G protein-coupled receptor
          of the glycoprotein hormone receptor subfamily
          Biochem. Biophys. Res. Commun. 247 (2), 266-270 (1998)
          98308104
          PUBMED 9642114
          2 (bases 1 to 2880)
          McDonald, T., Wang, R., Bailey, W. and Liu, Q.
          Direct Submission
          TITLE Submitted (29-APR-1998) Human Genetics, Merck Research Labs, West
          JOURNAL Point, PA 19486, USA
          FEATURES
            source
              1..2880
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /chromosome="12"
                /map="12q22-q23"
                49..2772
                /codon_start=1
                /product="orphan G protein-coupled receptor HG38"
                /protein_id="AAC28019.1"
                /db_xref="GI:3366802"
                /translation="MDTSRLGVLLSLPVLQLATGSSSPRSGLVLRGCPHCHCEPDG
                RMLKVDGSLGSLSPNLSVFTSYLDLNNISQLLPNPLPSLRFLELRAGNAL
                TYIPKAGTSLYKLVMLQNNQRLHVPTEALONLSQSLRLDANHSYVPPSCFSG
                LHSRLHLDNFALTEIPVQAFSLQALQAMTALNKTHIIPDYAFGNSLSVLVHLH
                NNRHSLGKCFDGLHSLETLNLNVDLDEFTALRTLSNLKELGFSNNRSIPEKA
                FVGPESLTIHFYDNPICOFVGRSAPHLPELRTLTNGASQITREPDLTGTANLESIT
                LTGAQISLPQVNCMLNQLVLDLSYNLLEDLPFSVCQKLQKIDLRHNIYEIKVD
                TFQQLLSLRSLNANWKIAIHPNAPSTLPSLIKLIDLSNLLSPFPIHGLHLHLKL
                TGNHALQSLISENPELVKVIEMPYAYQCCAFVCENAYKISNOWNKDNSSMDLHK
                KDAGMFQADRDLEDFLLDPEEDLKALHSVCSPSPGPKPKCBHLLDGLIRIGVMT
                IAVLALTCNALVTSTVRSPLYISPIKLLIGVIAAVNMLTGVSASVAVGADFTGSP
                ARHGAWNGVGVHIGLSIFASESSVPLTLAALERGFSVKYSAKPETKAPSSSLK
                VIIICALALUTMAVPLGLGSKYSGSEPLTLPPEPSTMGVYVALILLNLSLFLMM
                TIATYKLYCNLDKGLDENWPCSVKXHIALLFTNCILNCPVALFSLSLNLFTFSP
                EVIRFVLVVPDLACNLPLYILFNPHFKEDLSLRKQTYVTRSKHPSLMSINSD
                VEKQSCDSQALVITFTSSITYDLPPSSVPSVPVPTESCHLSSVAFVPL"
            CDS
              169 GAGCCGAGCGGAGGATGTTGCTAGGGTGGAATCTGCTCGAAGCTGGGGCTCTCGGAGCTG 228
              181 CCTTCCAACTCAGCGTCTTTCACTCTCTACCTCCTAGACCTCAGTATGAACAACATCACTCAG 240
              229 CTTTCCAACTCAGCGTCTTTCACTCTCTACCTCCTAGACCTCAGTATGAACAACATCACTCAG 288
              241 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGGAGTGTACGTTCTTGGGGAAC 300
              289 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGGAGTGTACGTTCTTGGGGAAC 348
              301 GCTCTGACATACATTCCTCAAGGAGCATTCACATGCGCTTTACAGTCTTAAAGTCTTTATG 360
              349 GCTCTGACATACATTCCTCAAGGAGCATTCACATGCGCTTTACAGTCTTAAAGTCTTTATG 408
              361 CTGCAGAAATATCAGCTAAGACACGTAACCAAGAGCTCTGCAGAAATTTGGGAAGCTTT 420
              409 CTGCAGAAATATCAGCTAAGACACGTAACCAAGAGCTCTGCAGAAATTTGGGAAGCTTT 468
              421 CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
              469 CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 528
              481 CTGCATTCCTGAGGACCTCTGGCTGATGACATGCTTAAACAGAAATCCCGCTCCAG 540
              529 CTGCATTCCTGAGGACCTCTGGCTGATGACATGCTTAAACAGAAATCCCGCTCCAG 588
              541 GCTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGGCCCTGAAACAAATATACAC 600
              589 GCTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGGCCCTGAAACAAATATACAC 648
              601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGATCTCATCTCCATAACAAT 660
              649 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGATCTCATCTCCATAACAAT 708
              661 AGATCCACTCCCTGGGGAAGAAATGCTTGTGAGGCTCCACAGCTAGAGACTTTAGAT 720
              709 AGATCCACTCCCTGGGGAAGAAATGCTTGTGAGGCTCCACAGCTAGAGACTTTAGAT 768
              721 TTAATAATACAATAACCTTGTGAAATTCCTCCACTGCAATAGGACACTCTCCAACTTAA 780
              769 TTAATAATACAATAACCTTGTGAAATTCCTCCACTGCAATAGGACACTCTCCAACTTAA 828
              781 GAATAGGATTTATAGCAACAATATCAGGTGATACCTCTGAGAAAGCAATTTGTAGGCAAC 840
              829 GAATAGGATTTATAGCAACAATATCAGGTGATACCTCTGAGAAAGCAATTTGTAGGCAAC 888
              841 CCTTCTCTTATTAACAATACATTTCTATGCAATCCATCCCAATTTGTTGGGAGATCTGCT 900
              889 CCTTCTCTTATTAACAATACATTTCTATGCAATCCATCCCAATTTGTTGGGAGATCTGCT 948
              901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 960
              949 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 1008
              961 TTTCTGATTTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
              1009 TTTCTGATTTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCAAGATC 1068
              1021 TCATCTCTCTCTCAACCGCTCTGCAATCAGTATCTTAATCTCCAAGTGTAGATCTGCT 1080
              1069 TCATCTCTCTCTCAACCGCTCTGCAATCAGTATCTTAATCTCCAAGTGTAGATCTGCT 1128
              1081 TACAACCTATTAGAAATTTTACCAAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC 1140
              1129 TACAACCTATTAGAAATTTTACCAAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC 1188
              1141 CTAAGACATAAATGAAATCTACGAAATTTAAAGTTGACATTTTCCAGAGTGTGTTAGCTTC 1200
              1189 CTAAGACATAAATGAAATCTACGAAATTTAAAGTTGACATTTTCCAGAGTGTGTTAGCTTC 1248
              1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
              1249 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1308

Db      169 GAGCCGAGCGGAGGATGTTGCTAGGGTGGAATCTGCTCGAAGCTGGGGCTCTCGGAGCTG 228
Db      181 CCTTCCAACTCAGCGTCTTTCACTCTCTACCTCCTAGACCTCAGTATGAACAACATCACTCAG 240
Db      229 CTTTCCAACTCAGCGTCTTTCACTCTCTACCTCCTAGACCTCAGTATGAACAACATCACTCAG 288
Db      241 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGGAGTGTACGTTCTTGGGGAAC 300
Db      289 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGGAGTGTACGTTCTTGGGGAAC 348
Db      301 GCTCTGACATACATTCCTCAAGGAGCATTCACATGCGCTTTACAGTCTTAAAGTCTTTATG 360
Db      349 GCTCTGACATACATTCCTCAAGGAGCATTCACATGCGCTTTACAGTCTTAAAGTCTTTATG 408
Db      361 CTGCAGAAATATCAGCTAAGACACGTAACCAAGAGCTCTGCAGAAATTTGGGAAGCTTT 420
Db      409 CTGCAGAAATATCAGCTAAGACACGTAACCAAGAGCTCTGCAGAAATTTGGGAAGCTTT 468
Db      421 CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db      469 CAATCCCTGCGTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 528
Db      481 CTGCATTCCTGAGGACCTCTGGCTGATGACATGCTTAAACAGAAATCCCGCTCCAG 540
Db      529 CTGCATTCCTGAGGACCTCTGGCTGATGACATGCTTAAACAGAAATCCCGCTCCAG 588
Db      541 GCTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGGCCCTGAAACAAATATACAC 600
Db      589 GCTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGGCCCTGAAACAAATATACAC 648
Db      601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGATCTCATCTCCATAACAAT 660
Db      649 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGATCTCATCTCCATAACAAT 708
Db      661 AGATCCACTCCCTGGGGAAGAAATGCTTGTGAGGCTCCACAGCTAGAGACTTTAGAT 720
Db      709 AGATCCACTCCCTGGGGAAGAAATGCTTGTGAGGCTCCACAGCTAGAGACTTTAGAT 768
Db      721 TTAATAATACAATAACCTTGTGAAATTCCTCCACTGCAATAGGACACTCTCCAACTTAA 780
Db      769 TTAATAATACAATAACCTTGTGAAATTCCTCCACTGCAATAGGACACTCTCCAACTTAA 828
Db      781 GAATAGGATTTATAGCAACAATATCAGGTGATACCTCTGAGAAAGCAATTTGTAGGCAAC 840
Db      829 GAATAGGATTTATAGCAACAATATCAGGTGATACCTCTGAGAAAGCAATTTGTAGGCAAC 888
Db      841 CCTTCTCTTATTAACAATACATTTCTATGCAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db      889 CCTTCTCTTATTAACAATACATTTCTATGCAATCCATCCCAATTTGTTGGGAGATCTGCT 948
Db      901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 960
Db      949 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 1008
Db      961 TTTCTGATTTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
Db      1009 TTTCTGATTTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCAAGATC 1068
Db      1021 TCATCTCTCTCTCAACCGCTCTGCAATCAGTATCTTAATCTCCAAGTGTAGATCTGCT 1080
Db      1069 TCATCTCTCTCTCAACCGCTCTGCAATCAGTATCTTAATCTCCAAGTGTAGATCTGCT 1128
Db      1081 TACAACCTATTAGAAATTTTACCAAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC 1140
Db      1129 TACAACCTATTAGAAATTTTACCAAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC 1188
Db      1141 CTAAGACATAAATGAAATCTACGAAATTTAAAGTTGACATTTTCCAGAGTGTGTTAGCTTC 1200
Db      1189 CTAAGACATAAATGAAATCTACGAAATTTAAAGTTGACATTTTCCAGAGTGTGTTAGCTTC 1248
Db      1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
Db      1249 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1308

Query Match 100.0%; Score 2724; DB 9; Length 2880;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACACTCCCGGCTCGGTGCTCTGCTGCTTGCCTTGGCTGTGCTGCTCAGCTGCGGACC 60
DB 49 ATGGACACTCCCGGCTCGGTGCTCTGCTGCTTGCCTTGGCTGTGCTGCTGAGCTGCGGACC 108
QY 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCACACACTGTCAATGTC 120
DB 109 GGGGGCAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCACACACTGTCAATGTC 168
QY 121 GAGCCGAGCGGAGGATGTTGCTAGGGTGGAATCTGCTCGAAGCTGGGGCTCTCGGAGCTG 180
```

Qy	1261	TTGGCATCCCTAATAAAAGCTGGACCTATCGTCCAACCCTCCTGTCGTCTTTTCCTATAACT	1320
Dd	1309	TTGCCATCCCTATAAAGCTGGACCTATCGTCCAACCCTCCTGTCGTCTTTTCCTATAACT	1368
Qy	1321	GGGTTACATGGTTTTAACTCACTTAATAAATAACAGGAATCATGCCTTTACAGAGCTTGATA	1380
Dd	1369	GGGTTACATGGTTTTAACTCACTTAATAAATAACAGGAATCATGCCTTTACAGAGCTTGATA	1428
Qy	1381	TCACTGAAAACCTTCCAGAACCTCMAGGTTATAGAAATGCTTATGCTTACACAGTGCCTGT	1440
Dd	1429	TCACTGAAAACCTTCCAGAACCTCMAGGTTATAGAAATGCTTATGCTTACACAGTGCCTGT	1488
Qy	1441	GCAATTGGAGTGTGAGAATGCCATAAGATTCTAATCAATGGGAATAAAGGTGACAAC	1500
Dd	1489	GCAATTGGAGTGTGAGAATGCCATAAGATTCTAATCAATGGGAATAAAGGTGACAAC	1548
Qy	1501	AGCAGTATGGAAGCACTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Dd	1549	AGCAGTATGGAAGCACTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG	1620
Dd	1609	GACCTTGAAGATTTCCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG	1668
Qy	1621	TGTTCACTTCCCAGAGCCCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC	1680
Dd	1669	TGTTCACTTCCCAGAGCCCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC	1728
Qy	1681	AGAAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Dd	1729	AGAAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1788
Qy	1741	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTGTAAATGGGGTCACTGCBA	1800
Dd	1789	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTGTAAATGGGGTCACTGCBA	1848
Qy	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGAATGGTTCACT	1860
Dd	1849	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGAATGGTTCACT	1908
Qy	1861	TTTTGGCAGCTTTGCAACACATGGTGCCTGGTGGGAGAAATGGGGTGGTTCGCCATGTCAIT	1920
Dd	1909	TTTTGGCAGCTTTGCAACACATGGTGCCTGGTGGGAGAAATGGGGTGGTTCGCCATGTCAIT	1968
Qy	1921	GGTTTTTTTGCCATTTTTTGGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGAGCCCTG	1980
Dd	1969	GGTTTTTTTGCCATTTTTTGGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGAGCCCTG	2028
Qy	1981	GAGCGTGGGTTCTCTGTGMAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040
Dd	2029	GAGCGTGGGTTCTCTGTGMAAATATTCTGCAAAATTTGNAACGAAAGCTCCATTTTCTAGC	2088
Qy	2041	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGCTGGCCCTTGACCAATGGCCGAGTCCCCCTG	2100
Dd	2089	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGCTGGCCCTTGACCAATGGCCGAGTCCCCCTG	2148
Qy	2101	CTGGGTGGCAGCAAGTATGGGGCCCTCCCTCTCTGCTGCTGCTTTGGCTTTTGGGGAGCCC	2160
Dd	2149	CTGGGTGGCAGCAAGTATGGGGCCCTCCCTCTCTGCTGCTGCTTTGGCTTTTGGGGAGCCC	2208
Qy	2161	AGCACCATGGGTACATGGTGCCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATGATG	2220
Dd	2209	AGCACCATGGGTACATGGTGCCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATGATG	2268
Qy	2221	ACCAATGGCTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG	2280
Dd	2269	ACCAATGGCTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG	2328
Qy	2281	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTGCTCTTTTCAACCACTGCATCTCTAACTGC	2340
Dd	2329	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTGCTCTTTTCAACCACTGCATCTCTAACTGC	2388

[illegible]

Db 2461 ATCTTGTTCAATCTCACTTTTAAGAGGATCTGGTGGAGCTGAGAAACCACTACGTC 2520
Qy 2521 TGGACAAGATCAAAACACCCAAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAACACAG 2580
Db 2521 TGGACAAGATCAAAACACCCAAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAACACAG 2580
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGTAACCTTTTACAGCTCCAGATCACTATATGACCTG 2640
Db 2581 TCCTGTGACTCAACTCAAGCCTTGTAACCTTTTACAGCTCCAGATCACTATATGACCTG 2640
Qy 2641 CCTCCAGTTCGGTCCCATCAACAGCTTATCCAGTACTGAGAGCTGCCATCTTCTCTCT 2700
Db 2641 CCTCCAGTTCGGTCCCATCAACAGCTTATCCAGTACTGAGAGCTGCCATCTTCTCTCT 2700
Qy 2701 GTGGCAATTTGCCATGTCTCTAA 2724
Db 2701 GTGGCAATTTGCCATGTCTCTAA 2724

RESULT 10
AF061444 2724 bp mRNA linear PRI 17-NOV-1998
LOCUS Homo sapiens G protein-coupled receptor LGR5 (LGR5) mRNA, complete cds.
DEFINITION
ACCESSION AF061444
VERSION AF061444.1 GI:3885471
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2724)
Hau, S.Y., Liang, S.G. and Heueh, A.J.
TITLE Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region
Mol. Endocrinol. 12 (12), 1830-1845 (1998)

JOURNAL MEDLINE
PUBMED 9849958
REFERENCE 2 (bases 1 to 2724)
Hau, S.Y., Liang, S.G. and Heueh, A.J.W.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (26-APR-1998) Gyn/Ob, Stanford University, 300 Pasteur Dr., Stanford, CA 95305-5317, USA

FEATURES
source
1..2724
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2724
/gene="LGR5"
1..2724
/gene="LGR5"
/note="Leucine-rich repeat containing G protein-coupled receptor; similar to glycoprotein hormone receptors; similar to FSH receptor, LH receptor and TSH receptor"
/codon_start=1
/product="G protein-coupled receptor LGR5"
/protein_id="AAC77911.1"
/db_xref="GI:3885472"

/translation="MDTSRLGVLLSLPVLQLATGGSSPSRSGVLLRGCPHCHCEPDG RMLLRVDCSLGLSLPSNLVSFTSYLDLSMNISQLPNPLSLHFLLEELRLAGNAL TVYPKAGFTGLYSLKVLMLNNQRLHVPTLEALQNLRLSLQSLDLNANHISVYPPSPFSG LNSRLHLDNDALEIIPVOAFRSLSALQMTALNKHITHIPDYAFGNLSMWVLHLH LNRHLSLKKCFDGLHLSLETLDLNNLDEFPARTLSNLKELGFPNSNNRSIPEKA FVGNPLSLTHIDNPIDFVGRSAFQHLPELRTLTNGASQITFFPDUTGTANLESUT LTGAQISLSPQVNCQNLQVLDLSYNLDELPSFVCQKQLKIDLRHNIYRIKVD TFQQLSLRSLNLANWKIAIHPNAPSTLPRLKIDLSNLLSPFITGLRHLYIKVD TGNHALQSLISSENPELVTEMVAYOCAGVCENAYKISNOWNKGDNSMDLHKW KDAGWFOQRDELDLEDLFEEDLKALHSVQCSPPGPKFCEHLLDGLIRIGWIT IAVLALNCALVTSTVFSRPLYSPIKLLIGVIAAVNMLTGVSSAVLAGVDAPFGSP ARHGAWENGVCVIGFSLIFASESSVFLTLAALERGFVSKYSAKFPETKAPFSSLK VIIICALLALTMAAVPLPGLGGSKYGASPLCLPLPFGEPSTWGYMVALILLNSLCPLMM

ORIGIN
Query Match 99.8%; Score 2719.2; DB 9; Length 2724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGACACCTCCCGGCTCGGTGTGCTCTCTTGGCCCTGTGCTCTGAGCTGGCGACC 60
Db 1 ATGACACCTCCCGGCTCGGTGTGCTCTCTTGGCCCTGTGCTCTGAGCTGGCGACC 60
Qy 61 GGGGGAGCTCTCCAGGTCTGGTGTGCTGAGGGGCTGCCCCACACACTGTCAATGC 120
Db 61 GGGGGAGCTCTCCAGGTCTGGTGTGCTGAGGGGCTGCCCCACACACTGTCAATGC 120
Qy 121 GAGCCGAGCGGAGGATGTTGCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Db 121 GAGCCGAGCGGAGGATGTTGCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Qy 181 CTTTCCAACTCAGCGCTCTTCACTCTTACCTTAGACCTCAGTATGAAACAATCAGTCAG 240
Db 181 CTTTCCAACTCAGCGCTCTTCACTCTTACCTTAGACCTCAGTATGAAACAATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTGCCCAGTCTCCACTTCTCTGGAGGAGTTACGTTCTTGGCGGAAAC 300
Db 241 CTGCTCCCGAATCCCTGCCCAGTCTCCACTTCTCTGGAGGAGTTACGTTCTTGGCGGAAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTTATG 360
Qy 361 CTGCAAGAATAATCAGTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGGCAAGCCTT 420
Db 361 CTGCAAGAATAATCAGTAAAGACAGTACCCACAGAGCTTTGCAGAAATTTGGCAAGCCTT 420
Qy 421 CAATCCCTGCTCTGGATGCTTAAACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
Db 421 CAATCCCTGCTCTGGATGCTTAAACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
Qy 481 CTGCATTTCCCTGGGACCTCTGGCTGGATGCAATGCTTAAAGAAATCCCGCTCCAG 540
Db 481 CTGCATTTCCCTGGGACCTCTGGCTGGATGCAATGCTTAAAGAAATCCCGCTCCAG 540
Qy 541 GCTTTAGAAAGTTTATCGGCATTCAGCCATGACCTTGGCCCTGAAACAAATACACAC 600
Db 541 GCTTTAGAAAGTTTATCGGCATTCAGCCATGACCTTGGCCCTGAAACAAATACACAC 600
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTCTACATCTCCATAAAT 660
Db 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTCTACATCTCCATAAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGAATGGGCTCCACAGCTAGAGACTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGAATGGGCTCCACAGCTAGAGACTTAGAT 720
Qy 721 TTAATTAACAATAACCTTGAATGAAATCCCACTGCAATTAGACACTCTCCAACTTAA 780
Db 721 TTAATTAACAATAACCTTGAATGAAATCCCACTGCAATTAGACACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTCTAGCAACAATATCAGGTCGATACCTTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTCTAGCAACAATATCAGGTCGATACCTTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATCTTCTAACAATCCCAATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATCAATTTCTAACAATCCCAATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACAATTTACCTGAACTAAGACACTGACTCTGAATGTGCTCACAATAACTGAA 960
Db 901 TTTCAACAATTTACCTGAACTAAGACACTGACTCTGAATGTGCTCACAATAACTGAA 960

QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db |||||
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db |||||
QY 1021 TCATCTCTTCTCAAACTGCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
Db |||||
QY 1021 TCATCTCTTCTCAAACTGCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
Db |||||
QY 1081 TACAACTTATAGAGATTTTACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db |||||
QY 1081 TACAACTTATAGAGATTTTACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db |||||
QY 1141 CTAAGACATAATGAAATCTAGAAATTAAGTGTGCACTTTCCAGCAGTGTCTTAGCCTC 1200
Db |||||
QY 1200 CGATCGCTGAATTTGGCTTGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db |||||
QY 1260 CGATCGCTGAATTTGGCTTGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db |||||
QY 1261 TTGCCATCCCTTAATAAGCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db |||||
QY 1320 TTGCCATCCCTTAATAAGCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db |||||
QY 1321 GGGTTACATGCTTAACTCACTTAAATTAACAGGAATCAATGCTTACAGAGCTTGATA 1380
Db |||||
QY 1380 GGGTTACATGCTTAACTCACTTAAATTAACAGGAATCAATGCTTACAGAGCTTGATA 1380
Db |||||
QY 1381 TCATCTGAAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGGCTGTGT 1440
Db |||||
QY 1440 TCATCTGAAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGGCTGTGT 1440
Db |||||
QY 1500 GCAATTTGAGTGTGAGAGTGCCTTAAGATTTTCAATGGAATGAATGAATGAATGAATGAAT 1500
Db |||||
QY 1501 GCAATTTGAGTGTGAGAGTGCCTTAAGATTTTCAATGGAATGAATGAATGAATGAATGAAT 1500
Db |||||
QY 1560 AGCAGTATGAGAGTGCCTTAAGATTTTCAATGGAATGAATGAATGAATGAATGAATGAAT 1560
Db |||||
QY 1561 GACCTTGAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db |||||
QY 1620 GACCTTGAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db |||||
QY 1621 TGTTCACTTCCCAAGGCTTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db |||||
QY 1681 TGTTCACTTCCCAAGGCTTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db |||||
QY 1681 AGAATTGAGTGTGAGACCATAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db |||||
QY 1741 AGAATTGAGTGTGAGACCATAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db |||||
QY 1741 ACAGTTTTTCCAGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db |||||
QY 1741 ACAGTTTTTCCAGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db |||||
QY 1801 CGAGTGAACATGCTCAACGAGTCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
QY 1801 CGAGTGAACATGCTCAACGAGTCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
QY 1861 TTTGGCAGCTTTGCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||
QY 1861 TTTGGCAGCTTTGCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCACTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db |||||
QY 1981 GGTTTTTTGTCCATTTTGTCTTCAAGATCACTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db |||||
QY 1981 GAGCTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAAGCAAGCTCCATTTTCTAGC 2040
Db |||||
QY 2041 CTGAAAGTAACTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100

Db 2041 CTGAAAGTAACTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 CTGGGTGGCAGCAAGTATGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db |||||
QY 2101 CTGGGTGGCAGCAAGTATGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db |||||
QY 2161 AGCACCATGGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db |||||
QY 2161 AGCACCATGGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db |||||
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGCAAGGAGGAGACCTGGGAAATATTTGG 2280
Db |||||
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGCAAGGAGGAGACCTGGGAAATATTTGG 2280
Db |||||
QY 2281 GACTGCTCTATGCTGTAACACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db |||||
QY 2281 GACTGCTCTATGCTGTAACACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db |||||
QY 2341 CCTGTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db |||||
QY 2341 CCTGTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db |||||
QY 2401 ATTAAGTTTATCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db |||||
QY 2401 ATTAAGTTTATCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db |||||
QY 2461 ATCTGTTCAATCTCTCTTTAAGGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db |||||
QY 2461 ATCTGTTCAATCTCTCTTTAAGGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db |||||
QY 2521 TGGCAAGATCAAAACACCAAGCTTGAATTTAATTAACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db |||||
QY 2521 TGGCAAGATCAAAACACCAAGCTTGAATTTAATTAACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db |||||
QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTTACCAGCTCCAGCATCCTTATGACCTG 2640
Db |||||
QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTTACCAGCTCCAGCATCCTTATGACCTG 2640
Db |||||
QY 2641 CCTCCCAGTTCGGGCAATCAACAGCTTATCCAGTGTGAGAGCTGAGAGCTGCTGCTGCTGCTGCT 2700
Db |||||
QY 2641 CCTCCCAGTTCGGGCAATCAACAGCTTATCCAGTGTGAGAGCTGAGAGCTGCTGCTGCTGCTGCT 2700
Db |||||
QY 2701 GTGGCATTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2724
Db |||||
QY 2701 GTGGCATTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2724
Db |||||

RESULT 11

BD135244
LOCUS
DEFINITION Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain.
ACCESSION
BD135244
VERSION
BD135244.1 GI:23230189
KEYWORDS
JP 2002507406-A/2.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2082)
AUTHORS
Hueh, A.J.W.; Heu, S.Y.; Liang, S.G. and Spe, P.J.V.D.
TITLE
Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain
JOURNAL
PATENT: JP 2002507406-A 2 12-MAR-2002;
THE BOARD OF TRUSTEES OF THE IELAND STANFORD JUNIOR UNIVERSITY,
AKZO NOBEL NV
COMMENT
OS Homo sapiens (human)
PN JP 2002507406-A/2
PD 12-MAR-2002
PF 25-MAR-1999 JP 2000537903
PR 26-MAR-1998 US 60/079501
PI AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES

	PI	VAN DER SPEK	C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19, C12N1/21, C12N5/10,C12P21/08,C12N15/00,C12N5/00 CC Novel mammalian G protein-coupled receptor having CC extracellular CC leucine-rich repeating domain FH key Location/Qualifiers FT source 1..2082 FT /organism='Homo sapiens (human)'. FEATURES source Location/Qualifiers 1..2082 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
	ORIGIN		
	Query Match	76.2%; Score 2074.6; DB 6; Length 2082;	
	Best Local Similarity	99.8%; Pred No. 0;	
	Matches 2077; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
Qy	643	CTACATCTCCATAACAATAGAACCTCACTCCCTGGGAAAGAATGCTTTGTATGGCTCCAC	702
Db	1	CTACATCTCCATAACAATAGAACCTCACTCCCTGGGAAAGAATGCTTTGTATGGCTCCAC	60
Qy	703	AGCCTAGAGACTTAGAATTAAATTACAATAACCTTCGATGAATGCCCACTGCNAATTAGG	762
Db	61	AGCCTAGAGACTTTAGATTTTAAATTAACAATAACCTTGATGAATGCCCACTGCAATTAGG	120
Qy	763	ACACTCTCCAACCTTAAAGAACTAGGATTTTCATAGCAACAATATCACGTCGATACCTGAG	822
Db	121	ACACTCTCCAACCTTAAAGAACTAGGATTTTCATAGCAACAATATCACGTCGATACCTGAG	180
Qy	823	AAAGCAATTTGTAGSCAACCCCTCTCTTTATTAACAATACATATTTCTATGCAATCCCATCCAA	882
Db	181	AAAGCAATTTGTAGSCAACCCCTCTCTTTATTAACAATACATATTTCTATGCAATCCCATCCAA	240
Qy	883	TTTTGTTGGGAGATCGCTTTTCAACATTTTACCCTGAACTAAGAACACTGACTCTGGAATGGT	942
Db	241	TTTTGTTGGGAGATCGCTTTTCAACATTTTACCCTGAACTAAGAACACTGACTCTGGAATGGT	300
Qy	943	GCCTCAAAAATAACTGAATTTCTTGATTTAACTGGAACTGCAAACCTGGAGAGCTTGACT	1002
Db	301	GCCTCAAAAATAACTGAATTTCTTGATTTAACTGGAACTGCAAACCTGGAGAGCTTGACT	360
Qy	1003	TTAACCTGGAGCAGAGATCTCATCTCTTCCTCAACACCGTCTGCAATCAAGTTACTTAATCTC	1062
Db	361	TTAACCTGGAGCAGAGATCTCATCTCTTCCTCAACACCGTCTGCAATCAAGTTACTTAATCTC	420
Qy	1063	CAAGTGTAGATCTGCTTCAACCTATTAGAGATTTACCCAGTGTTCCTCAGTCTGCCAA	1122
Db	421	CAAGTGTAGATCTGCTTCAACCTATTAGAGATTTACCCAGTGTTCCTCAGTCTGCCAA	480
Qy	1123	AAGCTTCAGAAAAATTGACCTAAGACATAAATGAATCTACGAAATTAAGTTGACACTTTTC	1182
Db	481	AAGCTTCAGAAAAATTGACCTAAGACATAAATGAATCTACGAAATTAAGTTGACACTTTTC	540
Qy	1183	CAGCAGTTGCTTAGCCCTCCGATCCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCAC	1242
Db	541	CAGCAGTTGCTTAGCCCTCCGATCCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTCAC	600
Qy	1243	CCCAATGCATTTTCCATCTTTGGCAATCCCTAATAAGCTGGACCTATCGTCCCAACCTCCTG	1302
Db	601	CCCAATGCATTTTCCATCTTTGGCAATCCCTAATAAGCTGGACCTATCGTCCCAACCTCCTG	660
Qy	1303	TCGTCCTTTTCCATAAATCGGGTTACATGTTTAACTCACTTTAAATTAACAGGAAATCAT	1362
Db	661	TCGTCCTTTTCCATAAATCGGGTTACATGTTTAACTCACTTTAAATTAACAGGAAATCAT	720
Qy	1363	GCCTTACAGAGCTTGATATCATCTGAAAACTTTCCAGAACTCAAGGHTATAGAAAATGCCT	1422
Db	721	GCCTTACAGAGCTTGATATCATCTGAAAACTTTCCAGAACTCAAGGHTATAGAAAATGCCT	780

Db 1861 AGAAGCAAACTAGCTGACAGAGTCAAAACACCCAGCTTGATGTCAATTAATCT 1920
Qy 2563 GATGATGCGAAACAGTCTCTGAGTCAACTCAAGCTTTGGTAACTTTACCAGTCC 2622
Db 1921 GATGATGCGAAACAGTCTCTGAGTCAACTCAAGCTTTGGTAACTTTACCAGTCC 1980
Qy 2623 AGCATCACTATGAGCTGCTCCAGTTCCTGCGCATCACAGCTTATCCAGTACTGAG 2682
Db 1981 AGCATCACTATGAGCTGCTCCAGTTCCTGCGCATCACAGCTTATCCAGTACTGAG 2040
Qy 2683 AGCTGCCATCTTCTCTGTGCGCATTTGTCCCATGCTCTCTAA 2724
Db 2041 AGCTGCCATCTTCTCTGTGCGCATTTGTCCCATGCTCTCTAA 2082

RESULT 12
AF110818 3115 bp mRNA linear ROD 11-FEB-1999
LOCUS Mus musculus orphan G protein-coupled receptor FEX mRNA, complete
DEFINITION cds.

ACCESSION AF110818

VERSION AF110818.1 GI:4262545

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 3115)

Identification of a novel seven-transmembrane receptor with

homology to glycoprotein receptors and its expression in the adult

and developing mouse

JOURNAL Biochem. Biophys. Res. Commun. 254 (1), 273-279 (1999)

MEDLINE 99121227

PUBMED 9920770

REFERENCE 2 (bases 1 to 3115)

AUTHORS Hermey, G., Methner, A., Schaller, H.C. and Hermans-Borgmeyer, I.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-1998) Center for Molecular Neurobiology,

Martinistr. 52, Hamburg D-20246, Germany

FEATURES Location/Qualifiers

source

1..3115

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1..2724

/note="contains leucine-rich repeats"

/codon_start=1

/product="orphan G protein-coupled receptor FEX"

/protein_id="AADI4684.1"

/db_xref="GI:4262546"

/translation="MDTSCVHMLLSLLALLQLVAAGSSPGSDAIPRGCPSHCHCELDG

RMLRVDCSLGLSELNLSNLSVFTSYLDLSMNNTSOLPASLHLRLCFLEELRLAGNAL

THIPKGAFTGLHSLKVLMLNNQKRVPEALQNLRSIQSLRLDANHSVYVPPSCFSG

LHSRLHLDNALIDTPVQAFSLALQAMTALNKIHLIADYAFGNLSLVVHLH

NNRIHSLGKCFDGLSEFLTNLNDLDEFFTAIKTSLNDELKELGFSNNIRSIPERA

FVGNPSLITHTFYDNPTQFVGSFAPQLPELRLTLINGASHITFPHLTGTATLESIT

LTGAKISLPQAVCDLPNLQVLDLSYLLDLPPLSGCQKLRIDLRHNEIYEIKGS

TFQOLFNLRLSNAPNFKIATHPNAPSPLSLIKLIDLSNLLSPFVTVGLHGLTHLKL

TGNALQSLIPSNAPPELKIETMPISAYVCCAFGCGENVYKISNOWNKDDGNSVDLHK

KDAGLFQVQDERDLEDFLLDEEDLNALHSVQSPSPGPKPCHEHLPGLSLIRIGWT

TAVLTSNCNALVPTVFTPLTISIKLLIGVIAVDILMGVSAVLAADVADFTGFR

AQHGAWEDGVCQIVGFLSIFASESSIFLLTLAALRGFSVKCSKFEKVPKPLFSR

AIVLVLVLLATITATPLLGSKYNAPLCLPLPFGEFTSTGYNVALVLLNSLCLFM

TIATYKLYCSLEKELENDWPCMKHIALLLFANCLYCPVAFLSFSLINLTFISF

DVIKFIILVIVPLPCLNPLLYIVFNPHFKEDMGLGKHTRFWMRSKHASLILNSDD

VEKRSQSTQALVSFTSHASIAVDLPSTSGSPAYPMTESCHLSVAVFVPCFL"

ORIGIN

Query Match

Best Local Similarity

Matches 2304; Conservative

75.4%; Score 2052.6; DB 10; Length 3115;

84.6%; Pred. No. 0;

Mismatches 419; Indels 0; Gaps 0;

Qy 1 ATGGACACCTCCCGGCTCGGTGTGCTCTCTGCTCCCTTGTGCTGTGCTGCTGACGCTGCGACG 60
Db 1 ATGGACACCTCCCTCGGCTCCATGCTCTCTGCTCTGCTGTGCTGCGGCTGCTGCTGCTGCGC 60
Qy 61 GGGGGCAGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCCCACACACTGTCAATGTC 120
Db 61 GCGGGAGCTCACCGGACAGATGCGATACCGGGGGCTGCCCATCACACTGTCACTGT 120
Qy 121 GAGCCCAAGGACAGATGTTGCTCAGGGTGAATGCTCCGACCTGGGGCTCTCGAGAGCTG 180
Db 121 GAGCTGATGGCAGATGCTGCTCAGGGTGAATGCTCCGACCTGGGGCTCTCGAGAGCTG 180
Qy 181 CCTTCGAACCTCAGCTCTTCACTCTACCTCAGTATGACCAACATCAGTCA 240
Db 181 CCTTCGAACCTCAGCTCTTCACTCTACCTCAGTATGACCAACATCAGTCA 240
Qy 241 CTGCTCCCGAATCCCTGCGGCTCTCGGCTTCTCGGAGGAGTTACGCTCTTGGCGGAAAC 300
Db 241 CTACCCGCGAGTCTCTACATCGCTCTGCTTCTAGAGAGTTACGCTCTTGGTGGAAAT 300
Qy 301 GCTCTGACATACATTCGAGGGAGCATTCACCTGGGCTTTACAGTCTTAAAGTCTTATG 360
Db 301 GCTTTGACACATTCGAGGGAGCGTTCAAGGGCTTCAAGGCTCAAGAGTCTTATG 360
Qy 361 CTGCAGATTAATCAGCTAAGACACGTAACCCACACAGCTCTGCAGAAATTTTCGGAAGCCTT 420
Db 361 CTGCAGAAACACAGCTGAGAAAGGTTCCGGAGGAGCGCTACAGAAATTTGAGAAAGCCTT 420
Qy 421 CAATCCCTGGGCTCGGATGCTTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 CAATCCCTGGGCTAGATGCCAACCAATCAGCTACGTGCCACCAGCTGTTTCAGGGGC 480
Qy 481 CTGCAATTCCTGAGGACCTGTGTGCTGATGACAAATGCGTTTAAAGAAATCCCCGTCAG 540
Db 481 CTGCACTCCCTGAGGACCTGTGTGCTAGTGAATGCTCTCACAGAGCTCCCTGTCTCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCATTCGAGGACCTGCGGCTGAGGAGGAGTTCGAGGAGGAGGAG 600
Db 541 GCTTTTGAAGTTTATCAGGCTTCGAGGAGGAGGAGTTCGAGGAGGAGGAGGAGGAGGAG 600
Qy 601 ATACGAGACTATGCTTTGGAAACCTCTCCAGCTGTTGTTTACATCTCCATAACAAT 660
Db 601 ATACGAGACTATGCTTTGGAAACCTCTCCAGCTGTTGTTTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAGAAAGAAATGCTTTGATGGGACTCCACAGCTAGAGACTTTAGAT 720
Qy 721 TTAATTTACATAACCTTGTGTAATCCCACTGCAATTAGGAGACTCTCCAACTTAA 780
Db 721 TTAATTTATATAACCTTGTGTAATCCCACTGCAATTAGGAGACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTCGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTCGAGGAGGAGGCTTCGAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGA CAATCCCATCAATTTGTTGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACATTTCTATGA CAACCCCATCAATTTGTTGGAGATCTGCT 900
Qy 901 TTTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCAACAATAACTGAA 960
Db 901 TTTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCAACAATAACTGAA 960
Qy 961 TTTTCTGATTTTAACTGGAACCTGCAAACTCGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Db 961 TTTTCTGATTTTAACTGGAACCTGCAAACTCGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTCTTCTCAAAAGCTGTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTCTTCTCAAAAGCTGTGATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT 1080

QY	1697	CCATAGCAGTTCTGGCAGCTTACTGTAAATGCTTTGGTGAATCTTCAACAGTTTTCAGATCCC	1756
Db	2078	CCATAGCAGTTCTGGCAGCTTACTGTAAATGCTTTGGTGAATCTTCAACAGTTTTCAGATCCC	2137
QY	1757	CTCTGTACATTTTCCCCATTAACCTGTTAAATGGGGTCATCGCAGCAGTGAACATGCTCA	1816
Db	2138	CTCTGTACATTTTCCCCATTAACCTGTTAAATGGGGTCATCGCAGCAGTGAACATGCTCA	2197
QY	1817	CGGAGTCTTCCAGTCCGCTGCTGGCTGGTGTGGATGCTTCTTTTGGCAGCTTTGGCAC	1876
Db	2198	CGGAGTCTTCCAGTCCGCTGCTGGCTGGTGTGGATGCTTCTTTTGGCAGCTTTGGCAC	2257
QY	1877	GACATGCTGCTGGTGGGAGATGGGGTGGT	1908
Db	2258	GACATGCTGCTGGTGGGAGATGGGGGAGAT	2289

AC078860 145165 bp. DNA linear PRI 23-JAN-2003
Homo sapiens 12 BAC RP11-186F10 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.

AC078860
AC078860.19 GI:13491193
HTG.

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 145165)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Banks, T., Barbacia, J.,
Benton, J., Bivaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D.,
Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,
Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J.,
Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W.,
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,
Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F.,
Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlesson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulsegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R.,
Martindale, A., Martinez, E., Massey, E., Mathew, E., McLeod, M.P.,
Meador, M., Mel, G., Merscher, S., Metzger, M., Miller, A., Miner, G.,
Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M.,
Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S.,
Ogulu, M., Okunolu, G., Orsugue, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C.,
Shooshkari, N., Siason, I., Sodergren, E., Sonaike, T., Sparks, A.,
Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A.,
Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, F., Telford, B.,
Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D.,
Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R.,
Washington, C., Watlington, S., Williams, G., Williamson, A.,
Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J.,
Zorilla, S., Zucherlapati, R. and Gibbs, R.

RESULT 14
AC078860/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

2 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (07-AUG-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (31-MAR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (01-APR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (23-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 2001 this sequence version replaced gi:13324671.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 145169

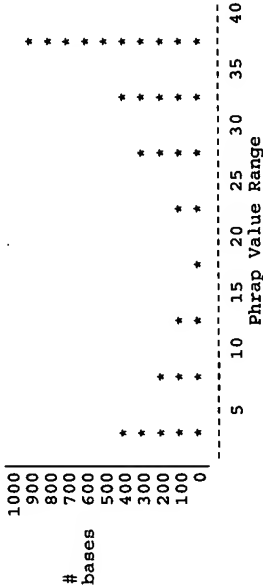
TITLE
Direct Submission
Unpublished

Phrap values in estimate:
Average error rate (BCM-Phrap estimate):
Fraction of Phrap values less than 40 :
Number of consensus changing edits:
Number of N's in consensus :

144455
5.28339e-05
0.0137759
19
0

Consensus changing edits		Edited+Context	
Position	Original+Context		
6882	agatgtggat (a)gctgtgctt	agatgtggat (g)gctgtgctt	
12828	tttcgggtca (n)tgcaactcc	tttcgggtca (c)tgcaactcc	
13151	aaaacaacat (n)caatagataa	aaaacaacat (a)caatagataa	
28957	ttgatattga (n)atggcatctt	ttgatattga (a)atggcatctt	
37139	atcgttatgt (n)tgcttaattt	atcgttatgt (c)tgcttaattt	
58030	ttttttattt (n)tatagttaat	ttttttattt (t)tatagttaat	
58290	tcattgacct (n)tcattgtccct	tcattgacct (a)tcattgtccct	
64729	ctatctctac (n)aaagaaaaga	ctatctctac (g)aaagaaaaga	
92472	aatggctttt (g)aaatgggccc	aatggctttt (a)aaatgggccc	
92662	cttagactgt (n)aatgtcagga	cttagactgt (g)aatgtcagga	
103773	gctaattttt (n)catttttagt	gctaattttt (a)catttttagt	
104056	cacatgcatt (n)ggatttttgc	cacatgcatt (a)ggatttttgc	
117534	tcaagcttgt (n)atttcagggt	tcaagcttgt (a)atttcagggt	
117581	cagatcatct (n)agggaatgct	cagatcatct (g)agggaatgct	
125485	gcttacacta (n)aaaatgaacc	gcttacacta (g)aaaatgaacc	
133723	agcattgtgt (n)nncacttaac	agcattgtgt (c)nncacttaac	
133724	gcacgggtgn (n)ncacttaata	gcacgggtgn (g)ncacttaata	
133725	catgggtgnn (n)cacttaataa	catgggtgnn (g)cacttaataa	
140801	aaatattgtc (n)taataaaagt	aaatattgtc (c)taataaaagt	

----- Distribution of Quality < 40 Bases -----



Version: 1.01 gxfp.
Location/Qualifiers
1. .145165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-186F10"
1. .2004
/note="overlaps bases 52259. .54262 of clone AC073612"
/function="clone overlap"
complement(6. .223)
/rpt_family="L1ME3"
repeat_region
complement(404. .698)
/rpt_family="L1M4"
repeat_region
complement(833. .1077)
/rpt_family="L1M4"
repeat_region
complement(1228. .1482)
/rpt_family="L1M4"
repeat_region
complement(1740. .2021)
/rpt_family="AluJb"
repeat_region
complement(2098. .2151)
/rpt_family="L1MEc"
repeat_region
complement(2152. .2449)
repeat_region

Qy	1334	TAAC	TGAC	TTAAAT	TTAAC	GGAAT	CA	TG	CTTAC	GAG	TTG	ATAT	CA	TG	AAACT	1339
Db	1334	TGAT	GCAC	CTG	AAAG	CTCAA	AGG	AACT	TTGG	CCCT	TGCT	TC	AG	GCCT	TTCT	1393
Qy	1394	TTCC	CAG	AACTCA	AG	TTTAT	AG	AAAT	TG	CTTAT	GC	TTAT	GC	TTAC	CTG	1453
Db	1394	TCC	CA	AACTG	AG	AT	CCT	GAG	GTG	CCCT	AC	CG	CTAC	CG	ATG	1453
Qy	1454	GTG	AG	AAATG	CC	TA	AG	ATTT	CT	AA	TG	GA	AT	TA	AG	1510
Db	1454	GTG	CC	AG	CTTCT	CA	AG	AC	CTCT	TGG	G	AG	CTG	GC	AG	1513
Qy	1511	ACG	AC	CTTCA	TA	AG	AA	AG	AT	GT	TT	CA	AG	CT	CA	1570
Db	1514	AGG	AG	GCAC	CA	AA	AG	AG	GC	CTG	GGT	CT	CT	GT	CG	1573
Qy	1571	ATTT	CT	CG	CTTG	AG	CTTT	TG	AG	-----	-	GA	AG	AC	CT	1621
Db	1574	TAG	AC	CTTGG	AT	AG	CT	TC	AG	AT	GG	GA	CA	G	AG	1633
Qy	1622	GTT	CA	CC	CTTCCC	AG	GC	CTTCAA	AC	CTCT	GT	GA	AC	CTCT	GT	1681
Db	1634	GC	AC	CTTGT	CA	AG	GC	CTTCA	AG	GC	CTCT	TT	GA	AG	CT	1693
Qy	1682	GA	AT	TG	AG	ATG	TG	CA	AT	AG	CTT	CT	GG	CA	CTT	1741
Db	1694	GC	CT	TG	CTGT	TG	GG	CC	AT	CG	TG	CT	CT	CG	TACT	1753
Qy	1742	CAG	TTT	-	-	-	-	-	-	-	-	-	-	-	-	1798
Db	1754	CAG	CT	TTTGG	C	AG	CG	GA	CC	AG	CC	GT	CT	CC	CC	1813
Qy	1799	CAG	CAG	TGA	CA	TG	CT	CA	CG	GG	AT	CT	CA	AG	TG	1858
Db	1814	CAG	CG	CC	CA	AG	CC	CT	GA	CG	GA	AT	TT	CT	CT	1873
Qy	1859	CTTT	TG	CAG	CTTT	GC	CA	CA	TG	CTG	TC	CT	GG	GA	AT	1918
Db	1874	CT	AT	TG	CT	CA	GT	TC	GG	AG	AG	CG	GT	CT	GG	1933
Qy	1919	TTG	GT	TTTT	TG	TC	AT	TTTT	GC	TT	CA	GA	AT	CT	GT	1978
Db	1934	CGG	CT	CT	CT	GG	CT	TC	AG	AG	CG	T	CG	GT	CT	1993
Qy	1979	TGG	AG	CG	TGG	TT	CT	GT	G	A	A	A	T	T	T	2038
Db	1994	TG	CAG	TG	CAG	AT	CT	CT	GT	AC	CT	CG	AG	CT	AG	2053
Qy	2039	GC	CT	G	A	A	G	T	A	A	T	T	G	A	A	2098
Db	2054	GC	GT	CC	GC	AG	GC	CA	CT	GG	AT	GC	CT	GG	GC	2113
Qy	2099	TG	CT	GG	GT	GG	CA	GA	AT	TG	GC	CT	CT	CT	CT	2152
Db	2114	TGG	CT	CG	GT	GG	GA	AT	TG	GC	CT	CT	CT	CT	CT	2173
Qy	2153	GGG	AG	CC	AG	CA	CA	CT	GG	CT	CA	CT	TT	GC	CT	2212
Db	2174	GC	GG	CC	GG	CC	CT	GG	CT	CT	CT	CT	CT	CT	CT	2233
Qy	2213	TC	AT	G	AT	G	A	T	T	G	C	T	A	T	T	2272
Db	2234	TG	GT	GG	TG	GG	CC	CT	AC	AT	CA	AG	CT	CT	AG	2293
Qy	2273	AT	AT	TG	GG	AG	CT	CT	AT	GG	CT	TT	CT	CT	CA	2332
Db	2294	CCG	GT	TG	GG	AG	CT	GG	CC	AT	GG	TG	CC	CA	CG	2353
Qy	2333	TAA	ACT	GC	CC	TG	GG	CT	TT	CT	GC	CT	TT	AA	T	2392
Db	2354	TCT	ACT	TG	CC	CG	GG	CT	TT	CT	CA	GT	TT	GG	CC	2413
Qy																

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 08:18:12 ; Search time 1573.87 Seconds
(without alignments)
10245.706 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 2724
Sequence: 1 atggacactcccgctcgg.....cattgtcccatgtctctaa 2724

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2724	100.0	2724	3 AAA30770	AAA30770 Human G p
2	2724	100.0	2724	8 AB242816	AB242816 Human G p
3	2724	100.0	2724	10 ADC22782	ADC22782 Human G p
4	2724	100.0	2724	10 ADH14255	ADH14255 Human HG3
5	2724	100.0	2724	12 ADI32985	ADI32985 Human G p
6	2724	100.0	2724	12 ADO29923	ADO29923 Human GPC
7	2724	100.0	2880	11 ADN33796	ADN33796 Cancer/an
8	2724	100.0	2880	13 ADO80249	ADO80249 G protein
9	2724	100.0	2373	13 ADP67869	ADP67869 Human HG3
10	2724	100.0	3032	11 ADN39627	ADN39627 Cancer/an
11	2724	100.0	3032	11 ADN39530	ADN39530 Cancer/an
12	2724	100.0	4558	2 AAX233980	AAX233980 Human HG3
13	2724	100.0	4570	12 ADL12472	ADL12472 Human ste
14	2720.8	99.9	2724	3 AAA30779	AAA30779 DNA encod
15	2720.8	99.9	2724	10 ADC22796	ADC22796 Human G p
16	2720.8	99.9	2724	10 ADH14269	ADH14269 Mutated h
17	2719.4	99.8	3438	10 ADF70582	ADF70582 Orphan re
18	2712	99.6	3297	2 AAX233981	AAX233981 Human HG3
19	2566	94.2	2651	10 ADB80463	ADB80463 Ovarian c
20	2566	94.2	2651	11 ADN39165	ADN39165 Cancer/an

21	2566	94.2	2651	11	ADN39795	ADN39795 Cancer/an
22	2074.6	76.2	2082	2	AA225344	AA225344 Human LGR
23	2052.6	75.4	2724	12	ADO30213	ADO30213 Mouse GPC
24	2052.6	75.4	3098	13	ADR67870	ADR67870 Mouse HG3
25	996	36.6	1790	6	ABQ55077	ABQ55077 Human ova
26	841.4	30.9	2901	6	AAI67921	AAI67921 Mouse LGR
27	841.4	30.9	2901	11	ADN02241	ADN02241 Mouse CDS
28	841.4	30.9	2901	12	ADK19406	ADK19406 ORF of CD
29	841.4	30.9	3637	6	AAI67920	AAI67920 Mouse LGR
30	841.4	30.9	3637	11	ADN02239	ADN02239 Mouse CDN
31	841.4	30.9	3637	12	ADK19404	ADK19404 cDNA enco
32	813.2	29.9	2901	6	AAI67927	AAI67927 Human LGR
33	813.2	29.9	2901	10	ADC16694	ADC16694 cDNA enco
34	813.2	29.9	2901	10	ADC16692	ADC16692 cDNA enco
35	813.2	29.9	2901	10	ADC16696	ADC16696 cDNA enco
36	813.2	29.9	2901	11	ADN02250	ADN02250 Human ful
37	813.2	29.9	2901	12	ADK19415	ADK19415 ORF of fu
38	813.2	29.9	3438	10	ABT31938	ABT31938 Human bre
39	813.2	29.9	3443	10	ADD89090	ADD89090 Encoding
40	813.2	29.9	3492	6	AAI67926	AAI67926 Human LGR
41	813.2	29.9	3492	11	ADN02248	ADN02248 Human ful
42	813.2	29.9	3492	12	ADK19413	ADK19413 Full leng
43	813.2	29.9	3492	12	ADQ15073	ADQ15073 Human can
44	813.2	29.9	3618	10	ADF70543	ADF70543 Orphan re
45	811.6	29.8	2901	10	ADC16695	ADC16695 cDNA with

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
XX
AC AAA30770;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.
XX
KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.
XX
OS Homo sapiens.
XX
PN WO200022129-A1.
XX
PD 20-APR-2000.
XX
PF 12-OCT-1999; 99WO-US023938.
XX
PR 13-OCT-1998; 98US-00170496.
XX
(AREN-) ARENA PHARM INC.
XX
Behan DP, Chalmers DT, Liaw CW;
XX
WPI; 2000-329165/28.
XX
P-PSDB; AAY90682.
XX
Non-endogenous constitutively activated human G protein-coupled
XX
receptors, useful for identifying agonists for use as pharmaceutical
XX
agents.
XX
Example 1; Page 315-317; 341pp; English.

The invention relates to constitutively active, non-endogenous versions of endogenous human orpan G protein-coupled receptors (GPCRs, AAY90643-CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743-CC and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

[illegible]

RESULT 2

RESULI 2
ABZ42816

ABZ42816
ID ABZ42816 standard: DNA: 2724 BP.

XX
ID ABZ42810

AC ABZ42816.

AC ABZ42816;
XX

XX
DE 04-MAR-2003 (E: yet contm.)DT 04-MAR-2003 (first entry)
yy[illegible]

Human G protein-coupled receptor GPR49 nucleotide SEQ ID NO:421.

G protein-coupled receptor; GPCR; antigenic peptide; Gene therapy;
G protein-coupled receptor modulator; antibody; immune-related disease;
growth-related disease; cell regeneration-related disease; AIDS; cancer;
immunological-related cell proliferative disease; autoimmune disease;
Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
ulcer; Gene; ds.

Homo sapiens.

WO200261087-A2.

08-AUG-2002.

19-DEC-2001; 2001WO-US050107.

19-DEC-2000; 2000US-0257144P.

(LIFE-) LIFESPAN BIOSCIENCES INC.

Burmer GC, Roush CL, Brown JP;

WPI: 2003-046718/04.

P-PSDB; ABP81968.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity and avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABP42523 to ABP42869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Query Match : 100.0%: Score 2724: DB 8: Length 2724:

Query Match :	100.0%;	Score 2/24;	DB 8;	Length 2/24;
Best Local Similarity	100.0%;	Pred	NO	0.

BEST LOCAL SIMILARITY 100.0%; Pred. No. 0;
 Matches 2724: Conservative 0: Mismatches 0: Indels 0:

Matches 2/24; Conservative 0; Mismatches 0; Indels 0;

1. **Abstract**

1 ATGGACACCTCCCGGTCGTCCTGTCCCTGCGCTGCGAGCTGGGACC 60

1 ATGGACACCTCCCGGCTCGGAGTGGCTGTCCCTGGCCGTGCTGCAGCTGGCGACC 60

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

b 1 ATGACACCTCCCGGCTCGGTGTGCTCCTGTCTCCTTGCCCTGTGCTGCAGCTGGCGACC 60

Db 61 GGGGCGAGCTCTCCAGGCTCTGGTGTGTCTGAGGGGCTGCCCCACACACTGTCTATGCG 120
Qy 121 GAGCCCGACGCGAGGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Db 121 GAGCCCGACGCGAGGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Qy 181 CCTTCCAACTCAGGGTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 181 CCTTCCAACTCAGGGTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTCCCTCCAGTCTCCGCTTCTGAGAGGATACGTCTTGGGGAAC 300
Db 241 CTGCTCCCGAATCCCTCCCTCCAGTCTCCGCTTCTGAGAGGATACGTCTTGGGGAAC 300
Qy 301 GCTCTGACATACATTTCCAAAGGAGCATCTACCTGGCCCTTTACAGTCTTTAAAGTCTTATG 360
Db 301 GCTCTGACATACATTTCCAAAGGAGCATCTACCTGGCCCTTTACAGTCTTTAAAGTCTTATG 360
Qy 361 CTGCAGATATCAGCTAAGACAGTACCCACAGAGCTCTGCAGATTTGGGAAGCCTT 420
Db 361 CTGCAGATATCAGCTAAGACAGTACCCACAGAGCTCTGCAGATTTGGGAAGCCTT 420
Qy 421 CAATCCCTCGCTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 480
Db 421 CAATCCCTCGCTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 480
Qy 481 CTGCATTTCCCTGAGGACCTGCTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTTCCCTGAGGACCTGCTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCAATGCGAATGCGAAGCCATGACCTTGGGCTCGAACAATAACAC 600
Db 541 GCTTTTGAAGTTTATCGGCAATGCGAAGCCATGACCTTGGGCTCGAACAATAACAC 600
Qy 601 ATACCACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGAGTCTTACATCTCCATAACAAT 660
Db 601 ATACCACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGAGTCTTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAATTAACAATTAACCTTGATGAATTTCCCACTGCAATTAGAGACACTCTCCAACTTAA 780
Db 721 TTAATTAACAATTAACCTTGATGAATTTCCCACTGCAATTAGAGACACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTCAAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTCAAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Db 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Qy 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Qy 1021 TCATCTCTTCTCAAACTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAACTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAGAGATTTACCAAGTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
Db 1081 TACAACCTATTAGAGATTTACCAAGTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
Qy 1141 CTAAGACATATGAATCTACGAAATTAAGTTGACATTTCCAGCAGTCTTAGCCTC 1200
Db 1141 CTAAGACATATGAATCTACGAAATTAAGTTGACATTTCCAGCAGTCTTAGCCTC 1200

Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTTATTCACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTTATTCACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCGCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db 1261 TTGCGCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440
Db 1381 TCATCTGAAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATTAAGGTTGCAAC 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATTAAGGTTGCAAC 1500
Qy 1501 AGCAGTATGAGACCTTTCATAAGAAAGATGCTCGAAATGTTTCAAGGCTCAAGATGAAAGT 1560
Db 1501 AGCAGTATGAGACCTTTCATAAGAAAGATGCTCGAAATGTTTCAAGGCTCAAGATGAAAGT 1560
Qy 1561 GACCTTGAAATTTCTGCTTGAATTTGAGGAAAGCCTGAAAGCCCTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAAATTTCTGCTTGAATTTGAGGAAAGCCTGAAAGCCCTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACCTTCCCGAGGCCCTTCAAACTGCTGAAACCTGCTGTAAGGCTGGCTGATC 1680
Db 1621 TGTTCACCTTCCCGAGGCCCTTCAAACTGCTGAAACCTGCTGTAAGGCTGGCTGATC 1680
Qy 1681 AGAATGAGTGTGGAACATAGCAGATTTCTGCACTTACTGTAATGCTTTGGTGAATTTCA 1740
Db 1681 AGAATGAGTGTGGAACATAGCAGATTTCTGCACTTACTGTAATGCTTTGGTGAATTTCA 1740
Qy 1741 ACAGTTTTCAGATCCCTCTGTATCAATTTCCCACTTAAATGTTAAATGGGGTCAATCGCA 1800
Db 1741 ACAGTTTTCAGATCCCTCTGTATCAATTTCCCACTTAAATGTTAAATGGGGTCAATCGCA 1800
Qy 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGCTGGCTGGTGGATGCGTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGCTGGCTGGTGGATGCGTCACT 1860
Qy 1861 TTTGGCAGCTTTGCAACACATGCTGCTGGTGGGAGAAATGGGGTGGTGGATGCGTCAAT 1920
Db 1861 TTTGGCAGCTTTGCAACACATGCTGCTGGTGGGAGAAATGGGGTGGTGGATGCGTCAAT 1920
Qy 1921 GGTTTTGTGTCATTTTGTGCTTCAAGATCAATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTGTGTCATTTTGTGCTTCAAGATCAATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAACTATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CTGAAAGTAACTATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 2101 CTGGGTGGCAGCAAGTATGGGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Qy 2161 AGCAACATGGGCTACATGGTGTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 AGCAACATGGGCTACATGGTGTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Qy 2221 ACCAATTCCTACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTGG 2280
Db 2221 ACCAATTCCTACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTGG 2280

Qy	361	CTGCAGAAATATCAGCTAAGACACGTTACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT	420
Db	361	CTGCAGAAATATCAGCTAAGACACGTTACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT	420
Qy	421	CAATCCCTCGGCTCTGGATGCTTAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Db	421	CAATCCCTCGGCTCTGGATGCTTAACCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Qy	481	CTGCAATTCCTCTAGGSCACCTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG	540
Db	481	CTGCAATTCCTCTAGGSCACCTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG	540
Qy	541	GCTTTTAAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGGAACAAAATACACCAAC	600
Db	541	GCTTTTAAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGGAACAAAATACACCAAC	600
Qy	601	ATACACAGCATGTGCTTTGGAACCTCTCGACGTTTGGTAGTTCTTACATCTCCATAAACAAT	660
Db	601	ATACACAGCATGTGCTTTGGAACCTCTCGACGTTTGGTAGTTCTTACATCTCCATAAACAAT	660
Qy	661	AGAAATCCATCCCTCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720
Db	661	AGAAATCCATCCCTCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT	720
Qy	721	TTAAATTTACAATAAACCCTTGATGAATTTCCCACTGCAATTAGCAACATCTCTCAACCTTAAA	780
Db	721	TTAAATTTACAATAAACCCTTGATGAATTTCCCACTGCAATTAGCAACATCTCTCAACCTTAAA	780
Qy	781	GAATCTAGGATTTCAATAGCAACAATATCAGCTGCGATCTCTGAGAAAGCAATTTGTAGGCAAC	840
Db	781	GAATCTAGGATTTCAATAGCAACAATATCAGCTGCGATCTCTGAGAAAGCAATTTGTAGGCAAC	840
Qy	841	CCCTCTCTTAATTAACAATPACTTTCTATGACAAATCCCATCCAAATTTGTTGGAGACTTGTCT	900
Db	841	CCCTCTCTTAATTAACAATPACTTTCTATGACAAATCCCATCCAAATTTGTTGGAGACTTGTCT	900
Qy	901	TTTTCACATTTACTGTAACCTTAAGAACACTGACTCTGTAATGGTGCCTCACAATAAATGAA	960
Db	901	TTTTCACATTTACTGTAACCTTAAGAACACTGACTCTGTAATGGTGCCTCACAATAAATGAA	960
Qy	961	TTTTCCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db	961	TTTTCCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Qy	1021	TCATCTCTTCTCTAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCT	1080
Db	1021	TCATCTCTTCTCTAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCT	1080
Qy	1081	TACAACTTATTAAGAAATTTACCCAGTTTTCAGTCTGCCAAAAGCTTTCAGAAAATGAC	1140
Db	1081	TACAACTTATTAAGAAATTTACCCAGTTTTCAGTCTGCCAAAAGCTTTCAGAAAATGAC	1140
Qy	1141	CTAAGACATTAATGAAATCTACGAAATTAAGTTTGAACACTTTTCCAGCAGTTTGTAGCCCTC	1200
Db	1141	CTAAGACATTAATGAAATCTACGAAATTAAGTTTGAACACTTTTCCAGCAGTTTGTAGCCCTC	1200
Qy	1201	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260
Db	1201	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260
Qy	1261	TTGCGCATCCCTTAATAAGCTGGACCTATCGTCCAAACCTCTGTGCTCTTTTCTATAACT	1320
Db	1261	TTGCGCATCCCTTAATAAGCTGGACCTATCGTCCAAACCTCTGTGCTCTTTTCTATAACT	1320
Qy	1321	GGGTTACATGGTTTTAACTCATTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1321	GGGTTACATGGTTTTAACTCATTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Qy	1381	TCATCTGAAAACCTTTCAGAACTCAAGGTTATAGAAATGCGTTATGCTTACCAAGTGTCTGT	1440
Db	1381	TCATCTGAAAACCTTTCAGAACTCAAGGTTATAGAAATGCGTTATGCTTACCAAGTGTCTGT	1440

QY	1441	GCATTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1500
DB	1441	GCATTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1500
QY	1501	AGCAGTATGGACGACCTTTCATAGAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT	1560
DB	1501	AGCAGTATGGACGACCTTTCATAGAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT	1560
QY	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
DB	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
QY	1621	TGTTCACTTCCCAGGCCCCCTTCAAAACCTGTGGAACACCTGCTTGATGGCTGGCTGATC	1680
DB	1621	TGTTCACTTCCCAGGCCCCCTTCAAAACCTGTGGAACACCTGCTTGATGGCTGGCTGATC	1680
QY	1681	AGAAATGGAGTGGACCATAGACAGTCTTGCGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
DB	1681	AGAAATGGAGTGGACCATAGACAGTCTTGCGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
QY	1741	ACAGTTTTAGATCCCCCTCTGTACATTTCCCCCAATTAATCTGTTAATTTGGGGTCATCGCA	1800
DB	1741	ACAGTTTTAGATCCCCCTCTGTACATTTCCCCCAATTAATCTGTTAATTTGGGGTCATCGCA	1800
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGGGAATGGTTCACT	1860
DB	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGGGAATGGTTCACT	1860
QY	1861	TTTGGGAGCTTTGCAAGACATGTGCTGTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT	1920
DB	1861	TTTGGGAGCTTTGCAAGACATGTGCTGTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT	1920
QY	1921	GGTTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG	1980
DB	1921	GGTTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG	1980
QY	1981	GAGCGTGGGTCTCTGTGAAATATCTGCAAAATTTGAAAACGAAAGCTCCATTTCTPAGC	2040
DB	1981	GAGCGTGGGTCTCTGTGAAATATCTGCAAAATTTGAAAACGAAAGCTCCATTTCTPAGC	2040
QY	2041	CTGAAAGTAATCATTTTGCTCTGTGCCCTGCTGGCCTTGACCATGGCCGAGTTCCCCCTG	2100
DB	2041	CTGAAAGTAATCATTTTGCTCTGTGCCCTGCTGGCCTTGACCATGGCCGAGTTCCCCCTG	2100
QY	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGCTTTTGGGAGCCCC	2160
DB	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGCTTTTGGGAGCCCC	2160
QY	2161	AGCACCATGGGCTACATGTCGCTCATCTGCTCAATTCCTTTGCTTCTCATGATG	2220
DB	2161	AGCACCATGGGCTACATGTCGCTCATCTGCTCAATTCCTTTGCTTCTCATGATG	2220
QY	2221	ACCATTCGCTACACCAAGCTCTACTGCAATTTTGGACAGGGAGACCTTGGAGAAATTTTGG	2280
DB	2221	ACCATTCGCTACACCAAGCTCTACTGCAATTTTGGACAGGGAGACCTTGGAGAAATTTTGG	2280
QY	2281	GACTGCTCTATGTTAAACACATTTGCCCTGCTCTTACCAACTGCACTCTTAAACTGC	2340
DB	2281	GACTGCTCTATGTTAAACACATTTGCCCTGCTCTTACCAACTGCACTCTTAAACTGC	2340
QY	2341	CCTGTGCTTTCTTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
DB	2341	CCTGTGCTTTCTTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
QY	2401	ATTAAGTTTATCTCTTGTGGTAGTCCACTTCTGTGCAATGTCATATCCCTTCTCTAC	2460
DB	2401	ATTAAGTTTATCTCTTGTGGTAGTCCACTTCTGTGCAATGTCATATCCCTTCTCTAC	2460
QY	2461	ATCTTGTTCATCTCTCATCTTTAAGGAGGATCTGGTGGCTGGAAGCAAACTACGTC	2520
DB	2461	ATCTTGTTCATCTCTCATCTTTAAGGAGGATCTGGTGGCTGGAAGCAAACTACGTC	2520
QY	2521	TGGACAAGATCAAAACACCCCAAGCTTGATGTCATTAACCTCTCATGATGTCGAAAACACAG	2580

Qy	721	TTAAATTACAATAACCTTGATGAATTC	CCCACTGCAATTAGGACACTCTCTCAA	CTTTAAA	780
Db	769	TTAAATTACAATAACCTTGATGAATTC	CCCACTGCAATTAGGACACTCTCTCAA	CTTTAAA	828
Qy	781	GAACTAGGATTTTCATAGCAAAATAT	CAGGTGCGATCTGAGAAAGCATTTGT	TAGGCAAC	840
Db	829	GAACTAGGATTTTCATAGCAAAATAT	CAGGTGCGATCTGAGAAAGCATTTGT	TAGGCAAC	888
Qy	841	CTTCTCTTTATTAACAATACATTTCT	ATGACAATCCCATCCAATTTGTTGGG	GAGATCTGCT	900
Db	889	CTTCTCTTTATTAACAATACATTTCT	ATGACAATCCCATCCAATTTGTTGGG	GAGATCTGCT	948
Qy	901	TTTCAACATTTACCTGAACTTAAGAA	CACTGACTCTGAAATGGTGGCTCACA	AAATAACTGAA	960
Db	949	TTTCAACATTTACCTGAACTTAAGAA	CACTGACTCTGAAATGGTGGCTCACA	AAATAACTGAA	1008
Qy	961	TTTCTCAATTTAACTGGAATCGCAAA	CCTGGAGAGTCTGACTTTTAACTGG	GAGACAGATC	1020
Db	1009	TTTCTCAATTTAACTGGAATCGCAAA	CCTGGAGAGTCTGACTTTTAACTGG	GAGACAGATC	1068
Qy	1021	TCATCTCTTCTCAAAACCGTCTGCAAT	CAGTTACGTTACTTAATTC	CAAGTGCTAGATCTGCT	1080
Db	1069	TCATCTCTTCTCAAAACCGTCTGCAAT	CAGTTACGTTACTTAATTC	CAAGTGCTAGATCTGCT	1128
Qy	1081	TACAACCTATTAGNAGATTTTACCCAG	GTTTTTACGCTGCGCAAAAGCTTCAGA	AAAAATTGAC	1140
Db	1129	TACAACCTATTAGNAGATTTTACCCAG	GTTTTTACGCTGCGCAAAAGCTTCAGA	AAAAATTGAC	1188
Qy	1141	CTAAGACATAATGAAATCTACGAAAT	TAAAGTTTGACACTTTTCCAGCAG	TGCTTAGTGCTC	1200
Db	1189	CTAAGACATAATGAAATCTACGAAAT	TAAAGTTTGACACTTTTCCAGCAG	TGCTTAGTGCTC	1248
Qy	1201	CGATCGCTGAAATTTGGCTTGGAA	CAAAATTTGCTATTATCACCCCAAT	GCAATTTTCCACT	1260
Db	1249	CGATCGCTGAAATTTGGCTTGGAA	CAAAATTTGCTATTATCACCCCAAT	GCAATTTTCCACT	1308
Qy	1261	TTGGCATCCCTTAATAAGCTTGGAC	CTATCGTCCAACTCTGCTGCTTTTCC	TATTAAC	1320
Db	1309	TTGGCATCCCTTAATAAGCTTGGAC	CTATCGTCCAACTCTGCTGCTTTTCC	TATTAAC	1368
Qy	1321	GGGTACATCGTTTAACTCACTTAA	ATAATTAACAGGAATCATGCTTTAC	AGAGCTTGATA	1380
Db	1369	GGGTACATCGTTTAACTCACTTAA	ATAATTAACAGGAATCATGCTTTAC	AGAGCTTGATA	1428
Qy	1381	TCATCTGAAAACCTTTCCAGAACT	CAAGGTTATAGAAATGCTTATGCTT	TACCAAGTGCTG	1440
Db	1429	TCATCTGAAAACCTTTCCAGAACT	CAAGGTTATAGAAATGCTTATGCTT	TACCAAGTGCTG	1488
Qy	1441	GCATTTGGAGTGTGAGATGCTTAAG	TATTTCTTAATCAATGGAATTAAG	GTGACAC	1500
Db	1489	GCATTTGGAGTGTGAGATGCTTAAG	TATTTCTTAATCAATGGAATTAAG	GTGACAC	1548
Qy	1501	AGCAGTATGGACGACCTTTCATA	GAAAGATGCTGGAATGTTT	CAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGACGACCTTTCATA	GAAAGATGCTGGAATGTTT	CAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCTGCTTGACTT	TGAGAGAGACCTGAAAGCCCTTCA	TTCAGTGCGAG	1620
Db	1609	GACCTTGAAGATTTCTGCTTGACTT	TGAGAGAGACCTGAAAGCCCTTCA	TTCAGTGCGAG	1668
Qy	1621	TGTTTCACTTTCCCAAGCCCTTCA	AAACCTGTGGAACACCTGCTGAT	GCTCGCTGATC	1680
Db	1669	TGTTTCACTTTCCCAAGCCCTTCA	AAACCTGTGGAACACCTGCTGAT	GCTCGCTGATC	1728
Qy	1681	AGAATTGGAGTGTGGACCATAGCAG	TGTTCTGGCACTTACTTGTAA	TGCTTTGGTGCACTTCA	1740
Db	1729	AGAATTGGAGTGTGGACCATAGCAG	TGTTCTGGCACTTACTTGTAA	TGCTTTGGTGCACTTCA	1788
Qy	1741	ACAGTTTTCAGATCCCTCTGTACAT	TTTCCCAATTAACATGTTAA	TGTTGGGTCATCGCA	1800
Db	1789	ACAGTTTTCAGATCCCTCTGTACAT	TTTCCCAATTAACATGTTAA	TGTTGGGTCATCGCA	1848

RESULT 8
ADQ80249
ID ADQ
XX
AC ADQ
XX
DT 21-
XX

Qy	1801	GCAGTGAACATGCTCACGGAGTCTCAGATGCCGTGCTGCTGGTGTGGATGGTTCACT	1861
Db	1849	GCAGTGAACATGCTCACGGAGTCTCAGATGCCGTGCTGCTGGTGTGGATGGTTCACT	1908
Qy	1861	TTTGGCAGCTTTGCAAGCATGCTGCTGGTGGAGAAATGGGGTTGGTTCGCCATGTCATT	1920
Db	1909	TTTGGCAGCTTTGCAAGCATGCTGCTGGTGGAGAAATGGGGTTGGTTCGCCATGTCATT	1968
Qy	1921	GGTTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	1969	GGTTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	2028
Qy	1981	GAGCGTGGGTTCTCTGTGAAATATTTCTGAAATAATTTGAAACGAAAGCTCAATTTTCTAGC	2040
Db	2029	GAGCGTGGGTTCTCTGTGAAATATTTCTGAAATAATTTGAAACGAAAGCTCAATTTTCTAGC	2088
Qy	2041	CTGAAAGTAATCATTTTTGTCTCTGCGCTCTGCGCTTGACCATGCGCGAGTTCCCCCTG	2100
Db	2089	CTGAAAGTAATCATTTTTGTCTCTGCGCTCTGCGCTTGACCATGCGCGAGTTCCCCCTG	2148
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGGGAGCCC	2160
Db	2149	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGGGAGCCC	2208
Qy	2161	AGCACCATGGGCTACATGGTGGCTCTCATCTTGCTCTGAATTCCTCTTGCTTCCTCATGATG	2220
Db	2209	AGCACCATGGGCTACATGGTGGCTCTCATCTTGCTCTGAATTCCTCTTGCTTCCTCATGATG	2268
Qy	2221	ACCAATGCCCTACACCAAGCTCTACTGCAATTTTCGACAAGGGAGACCTCGAGAGATATTTGG	2280
Db	2269	ACCAATGCCCTACACCAAGCTCTACTGCAATTTTCGACAAGGGAGACCTCGAGAGATATTTGG	2328
Qy	2281	GACTGCTCTATGGTAAACAACATTTGGCCCTTGCTCTTCCACCAACTGCATCTCTAAACTGC	2340
Db	2329	GACTGCTCTATGGTAAACAACATTTGGCCCTTGCTCTTCCACCAACTGCATCTCTAAACTGC	2388
Qy	2341	CCTGTGCTCTTCTTGCTCTCTCTCTTTTAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2400
Db	2389	CCTGTGCTCTTCTTGCTCTCTCTCTTTTAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2448
Qy	2401	ATTAAGTTTATCCTTCTGGTGGTAGTCCCACTTTCTGCAATGCTCAATCCCTTCTCTAC	2460
Db	2449	ATTAAGTTTATCCTTCTGGTGGTAGTCCCACTTTCTGCAATGCTCAATCCCTTCTCTAC	2508
Qy	2461	ATCTTGTTCATCCTCACTTTTAAGGAGGATCTGGTGAGCCTCGAGAAAGCAAAACCTACGTC	2520
Db	2509	ATCTTGTTCATCCTCACTTTTAAGGAGGATCTGGTGAGCCTCGAGAAAGCAAAACCTACGTC	2568
Qy	2521	TGGACAAGATCAAAAACCCCAAGCTTGATGTCATTTAACTCTGATGATGTGCGAAAACAG	2580
Db	2569	TGGACAAGATCAAAAACCCCAAGCTTGATGTCATTTAACTCTGATGATGTGCGAAAACAG	2628
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACCAAGCTCCAGCATCACTTATGACCTG	2640
Db	2629	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACCAAGCTCCAGCATCACTTATGACCTG	2688
Qy	2641	CCTCCCAAGTTCCGTGCCATCACCAAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTCT	2700
Db	2689	CCTCCCAAGTTCCGTGCCATCACCAAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTCT	2748
Qy	2701	GTGGCATTTGTGCCATGTCTCTAA	2724
Db	2749	GTGGCATTTGTGCCATGTCTCTAA	2772

Qy	241	CTGCTCCGAACTCCCTGCCAGTCTCGCTTCTGGAGGATTACGTCCTTGGCGGAAC	300
Db	441	CTGCTCCGAACTCCCTGCCAGTCTCGCTTCTGGAGGATTACGTCCTTGGCGGAAC	500
Qy	301	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTCTTTATG	360
Db	501	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTCTTTATG	560
Qy	361	CTGCAGAAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT	420
Db	561	CTGCAGAAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT	620
Qy	421	CAATCCCTGCGCTCGGATGCTAAACCATAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Db	621	CAATCCCTGCGCTCGGATGCTAAACCATAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	680
Qy	481	CTGCATTTCCCTGAGGCACCTGTGGCTGATGACAAATGGTTTAAACAGAAATCCCCTGCCAG	540
Db	681	CTGCATTTCCCTGAGGCACCTGTGGCTGATGACAAATGGTTTAAACAGAAATCCCCTGCCAG	740
Qy	541	GCTTTTGAAGTATTACGGCATTCGAAGCCATGACCTTGGCCCTGGAACAAATATCACCAAC	600
Db	741	GCTTTTGAAGTATTACGGCATTCGAAGCCATGACCTTGGCCCTGGAACAAATATCACCAAC	800
Qy	601	ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT	660
Db	801	ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT	860
Qy	661	AGAAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTTAGAT	720
Db	861	AGAAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTTAGAT	920
Qy	721	TTAAATTAACAATAACCTTGTGAATTTCCCACTGCAATTAGGACACTCTCCAACTTTAAA	780
Db	921	TTAAATTAACAATAACCTTGTGAATTTCCCACTGCAATTAGGACACTCTCCAACTTTAAA	980
Qy	781	GAACTAGGATTTATAGCAACAATATCAGTTCGATCTCTGAGGAAGCAATTTGTAGGCAAC	840
Db	981	GAACTAGGATTTATAGCAACAATATCAGTTCGATCTCTGAGGAAGCAATTTGTAGGCAAC	1040
Qy	841	CTTCTCTCTTATTAACAATACTTTCTATGACAAATCCCACTCAATTTGTTGGAGATCTGCT	900
Db	1041	CTTCTCTCTTATTAACAATACTTTCTATGACAAATCCCACTCAATTTGTTGGAGATCTGCT	1100
Qy	901	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATGAA	960
Db	1101	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAAATGAA	1160
Qy	961	TTTCTGTATTTTAACTGGNACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db	1161	TTTCTGTATTTTAACTGGNACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1220
Qy	1021	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTATCTCCAAGTCTAGATCTGTCT	1080
Db	1221	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTATCTCCAAGTCTAGATCTGTCT	1280
Qy	1081	TACAACTCTATTAGAGATTATCCCAAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC	1140
Db	1281	TACAACTCTATTAGAGATTATCCCAAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC	1340
Qy	1141	CTAAGACATATGAATCTAGCAAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC	1200
Db	1341	CTAAGACATATGAATCTAGCAAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC	1400
Qy	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT	1260
Db	1401	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT	1460
Qy	1261	TTGCCATCCCTTAATAAGCTGGAACCTATCGTCCAAACCTTCCTGTGCTCTTTTCTTATAACT	1320
Db	1461	TTGCCATCCCTTAATAAGCTGGAACCTATCGTCCAAACCTTCCTGTGCTCTTTTCTTATAACT	1520

Qy	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGAAATCATGCTTACAGAGCTTGATA	1380
Db	1521	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGAAATCATGCTTACAGAGCTTGATA	1580
Qy	1381	TCATCTGAAAACTTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACCAGTGTGT	1440
Db	1581	TCATCTGAAAACTTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACCAGTGTGT	1640
Qy	1441	GCATTTGGAGTGTGAGAAATGCCATATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1500
Db	1641	GCATTTGGAGTGTGAGAAATGCCATATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1700
Qy	1501	AGCAGTATGAGACACCTTTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT	1560
Db	1701	AGCAGTATGAGACACCTTTTCAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT	1760
Qy	1561	GACCTTGAAGATTTTCTGCTTGAATTTTGGAGAAAGACCTGAAAGCCCTTCATTTCAAGTGCAAG	1620
Db	1761	GACCTTGAAGATTTTCTGCTTGAATTTTGGAGAAAGACCTGAAAGCCCTTCATTTCAAGTGCAAG	1820
Qy	1621	TGTTTCACTTTCCCGAGGCCCTTCAAAACCCCTGTGAAACACCTGTGCTTGAATGGCTGTGATC	1680
Db	1821	TGTTTCACTTTCCCGAGGCCCTTCAAAACCCCTGTGAAACACCTGTGCTTGAATGGCTGTGATC	1880
Qy	1681	AGAAATCGAGTGTGGAACCATAGCAGATTTCTGACACTTACTTGTAAATGCTTTTGGTGACTTCA	1740
Db	1881	AGAAATCGAGTGTGGAACCATAGCAGATTTCTGACACTTACTTGTAAATGCTTTTGGTGACTTCA	1940
Qy	1741	ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCATTAACCTGTTTAAATTTGGGGTCAATGCA	1800
Db	1941	ACAGTTTTTCAGATCCCTCTGTGTAATTTTCCCCCATTAACCTGTTTAAATTTGGGGTCAATGCA	2000
Qy	1801	GCAGTGAACATGCTCAAGGAGATCTCCAGTGCCTGTGCTGGCTGTGCTGTGCTGTGCTCACT	1860
Db	2001	GCAGTGAACATGCTCAAGGAGATCTCCAGTGCCTGTGCTGGCTGTGCTGTGCTGTGCTCACT	2060
Qy	1861	TTTGGCAGCTTTTGCACACATATGGTGTGGGAGAAATGGGGTGTGGTGGCAATGCTCAAT	1920
Db	2061	TTTGGCAGCTTTTGCACACATATGGTGTGGGAGAAATGGGGTGTGGTGGCAATGCTCAAT	2120
Qy	1921	GGTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGTGGCAGCCCTG	1980
Db	2121	GGTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGTGGCAGCCCTG	2180
Qy	1981	GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGGAAGCTCCATTTTCTTAGC	2040
Db	2181	GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGGAAGCTCCATTTTCTTAGC	2240
Qy	2041	CTGAAAGTAAATCAATTTTGTCTGTGCCCTGTGCCCTTTGACCATGGCCGCAAGTTCCCTCTG	2100
Db	2241	CTGAAAGTAAATCAATTTTGTCTGTGCCCTGTGCCCTTTGACCATGGCCGCAAGTTCCCTCTG	2300
Qy	2101	CTGGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGGAGGCC	2160
Db	2301	CTGGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCTGCTGCTTTTGGGGAGGCC	2360
Qy	2161	AGCACCATGGGCTACATGGTGGCTCTCATCTTGTGCTCAATTTCCCTTGTGCTCTCATGATG	2220
Db	2361	AGCACCATGGGCTACATGGTGGCTCTCATCTTGTGCTCAATTTCCCTTGTGCTCTCATGATG	2420
Qy	2221	ACCAATGCTTACACCAAGCTCTACTGCAATTTGGAACAAGGAGACCTGGAGAAATTTTGG	2280
Db	2421	ACCAATGCTTACACCAAGCTCTACTGCAATTTGGAACAAGGAGACCTGGAGAAATTTTGG	2480
Qy	2281	GACTGCTCTATGGTAAACACATTTGCCCTGTGCTTCTTACCAAACCTGATCCTTAAACCTG	2340
Db	2481	GACTGCTCTATGGTAAACACATTTGCCCTGTGCTTCTTACCAAACCTGATCCTTAAACCTG	2540
Qy	2341	CCTGTGCTTCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2400
Db	2541	CCTGTGCTTCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2600
Qy	2401	ATTAAGTTTATCTTCTGCTGCTAGTCCCACTTCTCTGATGCTCTCAATFCCCTCTCTTAC	2460

[illegible]

QY	1021	TCATCTCTTCTCAAAACCGTGCGAATCAGTTACTTAATCTCAAGTGTGATGCTGTCT	1081
DB	1282	TCATCTCTTCTCAAAACCGTGCGAATCAGTTACTTAATCTCAAGTGTGATGCTGTCT	1341
QY	1081	TACAACCTATTAGAAGATTTACCAGAGTTTTTCCAGTCTGCCAARAAGCTTCAGAAAATTGC	1140
DB	1342	TACAACCTATTAGAAGATTTACCAGAGTTTTTCCAGTCTGCCAARAAGCTTCAGAAAATTGC	1401
QY	1141	CTAAGACATAATGAANAATTACGAANAATTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC	1200
DB	1402	CTAAGACATAATGAANAATTACGAANAATTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC	1461
QY	1201	CGATCGCTGAATTTGGCTTGGAAACAAANAATTGCTATTATTACCCCCAATGCATTTTCCACT	1260
DB	1462	CGATCGCTGAATTTGGCTTGGAAACAAANAATTGCTATTATTACCCCCAATGCATTTTCCACT	1521
QY	1261	TTGGCATCCCTAATAAGCTGGACCTTATCGTCCAACCTCTGTCTGTCTTTTCTCTATAAAT	1320
DB	1522	TTGGCATCCCTAATAAGCTGGACCTTATCGTCCAACCTCTGTCTGTCTTTTCTCTATAAAT	1581
QY	1321	GGGTTACATGGTTTTAACTCATTTAAANAATTAAACAGGAATCATGCCTTTACAGAGCTTGATA	1380
DB	1582	GGGTTACATGGTTTTAACTCATTTAAANAATTAAACAGGAATCATGCCTTTACAGAGCTTGATA	1641
QY	1381	TCATCTGAAAACTTTCCAGAACTCAAGGTTATTAGAAATGCTTTATGCTTACCAGTGTCTGT	1440
DB	1642	TCATCTGAAAACTTTCCAGAACTCAAGGTTATTAGAAATGCTTTATGCTTACCAGTGTCTGT	1701
QY	1441	GCATTTGGAGTGTGAGATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAC	1500
DB	1702	GCATTTGGAGTGTGAGATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAC	1761
QY	1501	AGCAGTATGACGACCTTTCTTAAGAAAAGATGCTTGGAAATGTTTCAAGGCTCAAGAATGAACGT	1560
DB	1762	AGCAGTATGACGACCTTTCTTAAGAAAAGATGCTTGGAAATGTTTCAAGGCTCAAGAATGAACGT	1821
QY	1561	GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGACG	1620
DB	1822	GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGACG	1881
QY	1621	TGTTCACTTTCCCAGGCCCTTTCAAACCTGTGAAACACTGCTTGATGGCTGGCTGATC	1680
DB	1882	TGTTCACTTTCCCAGGCCCTTTCAAACCTGTGAAACACTGCTTGATGGCTGGCTGATC	1941
QY	1681	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCTCA	1740
DB	1942	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCTCA	2001
QY	1741	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTCAATCGCA	1800
DB	2002	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTCAATCGCA	2061
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTCTGGCTGGTGTGATGCGTTCACT	1860
DB	2062	GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTCTGGCTGGTGTGATGCGTTCACT	2121
QY	1861	TTTGGCAGCTTTGCACGACATGCTGCTGTGGGAGAAATGGGGTGTGGTTGCCATGTCAATT	1920
DB	2122	TTTGGCAGCTTTGCACGACATGCTGCTGTGGGAGAAATGGGGTGTGGTTGCCATGTCAATT	2181
QY	1921	GGTTTTTGTGCTTTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	1980
DB	2182	GGTTTTTGTGCTTTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	2241
QY	1981	GAGGTGGGTCTCTGTGAAATATTCTGCNAANAATTTGAAAACGAAAGCTCCATTTTCTAGC	2040
DB	2242	GAGGTGGGTCTCTGTGAAATATTCTGCNAANAATTTGAAAACGAAAGCTCCATTTTCTAGC	2301
QY	2041	CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGGCTTGCACCATGSCCGCAGTTCCCTCTG	2100
DB	2302	CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGGCTTGCACCATGSCCGCAGTTCCCTCTG	2361
QY	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTCTTTTGTGCTTTTGGGGAGCC	2160

Db 481 CTGCATTCCTGAGGCACCTGTGGCTGGATGACAAATGCGTTAAACGAAATCCCCTGCAG 540
Qy 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAATAATCAATAACCTTTGATGAATTCCTCCACTGCAATTTAGGACACTCTCAACCTTAAA 780
Db 721 TTAATAATCAATAACCTTTGATGAATTCCTCCACTGCAATTTAGGACACTCTCAACCTTAAA 780
Qy 781 GAACTAGGATTTCTAGCAACAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAC 840
Db 781 GAACTAGGATTTCTAGCAACAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAC 840
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Qy 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCCTCAAAATACGTAA 960
Db 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCCTCAAAATACGTAA 960
Qy 961 TTTCTGATTTTAACTGGAACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Db 961 TTTCTGATTTTAACTGGAACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAACTGCTAGATCTGCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAACTGCTAGATCTGCT 1080
Qy 1081 TACAACCTATTAGAGATTTACCCAGTTTTTTCAGTCTGCAAAAGCTTTAGAAAATTTGAC 1140
Db 1081 TACAACCTATTAGAGATTTTACCCAGTTTTTTCAGTCTGCAAAAGCTTTAGAAAATTTGAC 1140
Qy 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGAGAGTCTGCTAGCCTC 1200
Db 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGAGAGTCTGCTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTCACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTCACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCCCTATAACT 1320
Db 1261 TTGCCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGGCTGT 1440
Db 1381 TCATCTGAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGGCTGT 1440
Qy 1441 GCAATTTGGAGTGTGTGAGAACTGCTTATAGAAATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 1441 GCAATTTGGAGTGTGTGAGAACTGCTTATAGAAATTTCTAATCAATGGAATAAAGGTGACAA 1500
Qy 1501 AGCAGTATGAGAGCTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGAGAGCTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGAAAGCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGAAAGCTGAAAGCCCTTCATTCAGTGCAG 1620

Qy 1621 TGTTCACTTCCCAGAGGCCCTTCAAACCTGTGTGAAACACCTGTGATGGCTGGCTGATC 1680
Db 1621 TGTTCACTTCCCAGAGGCCCTTCAAACCTGTGTGAAACACCTGTGATGGCTGGCTGATC 1680
Qy 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGGCACTTACTTCTAATGCTTTTGGTGAATCA 1740
Db 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGGCACTTACTTCTAATGCTTTTGGTGAATCA 1740
Qy 1741 ACAGTTTTTCAGATCCCCTCTGTACATTTCCCCCAATTTAACTGTTAAATTTGGGGTCAAT 1800
Db 1741 ACAGTTTTTCAGATCCCCTCTGTACATTTCCCCCAATTTAACTGTTAAATTTGGGGTCAAT 1800
Qy 1801 GCAGTGAACATGCTCAAGGGAGTCTCCAGTCCCGTGTGCTGGTGTGGATGCGTTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGGAGTCTCCAGTCCCGTGTGCTGGTGTGGATGCGTTCACT 1860
Qy 1861 TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTGCATGTCAAT 1920
Db 1861 TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTGCATGTCAAT 1920
Qy 1921 GGTTTTTTGTCCATTTTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTTCCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTTCCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAATGATCATTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CTGAAATGATCATTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Qy 2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db 2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Qy 2221 ACCATTCCTTACACCAAGCTCTACTGCAATTTGCAAGGGAGACCTGGAGAAATTTTGG 2280
Db 2221 ACCATTCCTTACACCAAGCTCTACTGCAATTTGCAAGGGAGACCTGGAGAAATTTTGG 2280
Qy 2281 GACTGCTTATGTAAGTAAACACATTTGCTGCTCTTACCAACTGCTGCTTCTTAACTGTC 2340
Db 2281 GACTGCTTATGTAAGTAAACACATTTGCTGCTCTTACCAACTGCTGCTTCTTAACTGTC 2340
Qy 2341 CCTGTGGCTTTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db 2341 CCTGTGGCTTTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Qy 2401 ATTAAGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
Db 2401 ATTAAGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
Qy 2461 ATCTTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGGAGCTGAGAAAGCAAACTTAC 2520
Db 2461 ATCTTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGGAGCTGAGAAAGCAAACTTAC 2520
Qy 2521 TGGACAAAGATCAAAACCCCAAGCTTGAATGCTCAATTTAACTCTGTGATGTGTCGAAAA 2580
Db 2521 TGGACAAAGATCAAAACCCCAAGCTTGAATGCTCAATTTAACTCTGTGATGTGTCGAAAA 2580
Qy 2581 TCCTGTGATCACTCAAGCTTGGTAACTTTACAGCTCCAGCATCATTTATGACCTG 2640
Db 2581 TCCTGTGATCACTCAAGCTTGGTAACTTTACAGCTCCAGCATCATTTATGACCTG 2640
Qy 2641 CCTCCAGTTCCTGGCCATCAGCAGCTTATCTCAGTGAAGCTGAGCTGCTTCTTCTCTCT 2700
Db 2641 CCTCCAGTTCCTGGCCATCAGCAGCTTATCTCAGTGAAGCTGAGCTGCTTCTTCTCTCT 2700

QY	2701	GTGGCATTGTCCCATGTCTCTAA	2724	Db	1	ATGGACACCTCCCGGCTCGGTGCTGCTCTGCTTGTGCTGTGCTGTGCTGCGAGC	60
Db	2701		2724	QY	61	GGGGGAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATGC	120
				Db	61	GGGGGAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATGC	120
RESULT 15				QY	121	GAGCCGACGCGGAGTGTGCTCAGGCTGAGCTCTCCGACCTGGGGCTCTCGGAGCTG	180
ADC22796				Db	121	GAGCCGACGCGGAGTGTGCTCAGGCTGAGCTCTCCGACCTGGGGCTCTCGGAGCTG	180
ID	ADC22796	standard; cDNA; 2724 BP.		QY	181	CTTTCCAACTCAGCGTCTTCACTCTACCTAGACCTCAGTATGAACAAATCAGTCAG	240
XX	AC	ADC22796;		Db	181	CTTTCCAACTCAGCGTCTTCACTCTACCTAGACCTCAGTATGAACAAATCAGTCAG	240
XX	18-DEC-2003	(first entry)		QY	241	CTGCTCCGGAATCCCTGGCCAGTCTCCGCTTCTCGGAGGAGTTACGTTTGGGGAAAC	300
DE	Human G protein-coupled receptor cDNA #79.			Db	241	CTGCTCCGGAATCCCTGGCCAGTCTCCGCTTCTCGGAGGAGTTACGTTTGGGGAAAC	300
XX	Human; gene; ss; G protein-coupled receptor; GPCR;			QY	301	GCTCTGACATACATTTCCCAAGGAGCAATCACTGGCCCTTTTACAGTCTTAAAGTTCTTATG	360
KW	transmembrane-6 region; TM6; intracellular-3 region; IC3.			Db	301	GCTCTGACATACATTTCCCAAGGAGCAATCACTGGCCCTTTTACAGTCTTAAAGTTCTTATG	360
XX	Homo sapiens.			QY	361	CTGCAGAAATATCAGCTAAGACACGTACCCACAGAAGCTCTGCAGAAATTTGCGAAGCCTT	420
OS	US655339-B1.			Db	361	CTGCAGAAATATCAGCTAAGACACGTACCCACAGAAGCTCTGCAGAAATTTGCGAAGCCTT	420
FN	29-APR-2003.			QY	421	CAATCCCTGGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
XX	13-OCT-1998; 98US-00170496.			Db	421	CAATCCCTGGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
XX	14-APR-1997; 97US-00839449.			QY	481	CTGCATTTCCCTGAGGACCTGTGGCTGATGACAAATGCTTTAAACAGAAATCCCGTCCAG	540
PR	14-APR-1998; 98US-00060188.			Db	481	CTGCATTTCCCTGAGGACCTGTGGCTGATGACAAATGCTTTAAACAGAAATCCCGTCCAG	540
PR	26-JUN-1998; 98US-0090783P.			QY	541	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC	600
PR	07-AUG-1998; 98US-0095677P.			Db	541	GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC	600
PA	(AREN-) ARENA PHARM INC.			QY	601	ATACGACACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTCTACATCTCCATAACAT	660
XX	Liaw CW, Behan DP, Chalmers DT;			Db	601	ATACGACACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTCTACATCTCCATAACAT	660
XX	WPI; 2003-742861/70.			QY	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
PT	P-PSDB; ADC22797.			Db	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
XX	Creating a constitutively active version of an endogenous human G protein			QY	721	TTAAATTAACAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA	780
CC	coupled receptor (GPCR) comprises substituting a specific amino acid in			Db	721	TTAAATTAACAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA	780
CC	the transmembrane-6 region with a different amino acid, and testing for			QY	781	GAACTAGGATTTCAATAGCAACAAATATCAGTTCGATCTGAGTGAAGCAATTTGTAGGCAAC	840
CC	constitutive activity.			Db	781	GAACTAGGATTTCAATAGCAACAAATATCAGTTCGATCTGAGTGAAGCAATTTGTAGGCAAC	840
CC	Example 2; SEQ ID NO 277; 221pp; English.			QY	841	CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT	900
CC	The invention relates to a method for treating a non-endogenous,			Db	841	CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT	900
CC	constitutively active version of an endogenous human G protein-coupled			QY	901	TTTCAACATTTACTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA	960
CC	receptor (GPCR) that has a transmembrane-6 (TM6) region and an			Db	901	TTTCAACATTTACTGAACTAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA	960
CC	intracellular-3 (IC3) region, by substituting a specific amino acid in			QY	961	TTTCTGATTTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
CC	the TM6 region with a different amino acid, and testing for constitutive			Db	961	TTTCTGATTTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
CC	activity. The method is useful for creating a constitutively active			QY	1021	TCATCTCTTCTCTCAACCGTCTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT	1080
CC	version of an endogenous human GPCR that comprises a transmembrane 6			Db	1021	TCATCTCTTCTCTCAACCGTCTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT	1080
CC	region and an intracellular loop 3 region. The altered human GPCR			QY	1081	TACAACTATTAGAAAGTTTACCCAGTTTTTTCAGTCTGCGCCAAAAGCTTTCAGAAAATTGAC	1140
CC	polypeptides are useful for screening test compounds for identification			Db	1081	TACAACTATTAGAAAGTTTACCCAGTTTTTTCAGTCTGCGCCAAAAGCTTTCAGAAAATTGAC	1140
CC	of inverse agonists or partial agonists of GPCR polypeptides, which may						
CC	have therapeutic uses. The altered GPCRs may also be used in vivo or in						
CC	vitro in biological research. A nucleic acid encoding the altered GPCR						
CC	may be used to create a transgenic animal expressing the altered GPCR.						
CC	The method allows screening for compounds that modulate the activity of a						
CC	human G protein-coupled receptor without the need for provision of a						
CC	ligand for the receptor. This is particularly useful in allowing						
CC	screening of compounds against orphan receptors for which no ligand is						
CC	currently known. This sequence represents cDNA encoding a human GPCR						
CC	polypeptide of the invention.						
XX	Sequence 2724 BP; 667 A; 723 C; 553 G; 781 T; 0 U; 0 Other;						
XX	Query Match 99.9%; Score 2720.8; DB 10; Length 2724;						
XX	Best Local Similarity 99.9%; Pred. No. 0;						
XX	Matches 2722; Conservative 2; Mismatches 0; Gaps 0;						
QY	1	ATGGACACCTCCCGGCTCGGTGCTGCTCTGCTTGTGCTGTGCTGTGCTGCGAGC	60				

QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACCTTCCAGCAGTTGCTTAGCCTC 1200
DB 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACCTTCCAGCAGTTGCTTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTTACCCCAATGCAATTTCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTTACCCCAATGCAATTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTTCCCTATAACT 1320
DB 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGGTTAACTCACTTAAATTTAAACAGGAATCATGCTTCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGGTTAACTCACTTAAATTTAAACAGGAATCATGCTTCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAATCTTTCCAGACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTCTGT 1440
DB 1381 TCATCTGAAATCTTTCCAGACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTCTGT 1440
QY 1441 GCATTTGGAGTGTGAGAAATGCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
DB 1441 GCATTTGGAGTGTGAGAAATGCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
QY 1501 AGCAGTATGGACGACCTTCATAGAAAGATGCTGGAATGTTTCAAGCTCAAGATGAAGCT 1560
DB 1501 AGCAGTATGGACGACCTTCATAGAAAGATGCTGGAATGTTTCAAGCTCAAGATGAAGCT 1560
QY 1561 GACCTTGAAGATTTCTGCTTACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1561 GACCTTGAAGATTTCTGCTTACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCTCTGAAACACCTCTTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCTCTGAAACACCTCTTGATGGCTGGCTGATC 1680
QY 1681 AGAATTGAGTGTGACCAATAGCAGTTCTGGCACTTACCTGTAATGCTTGGTGAATGCTCA 1740
DB 1681 AGAATTGAGTGTGACCAATAGCAGTTCTGGCACTTACCTGTAATGCTTGGTGAATGCTCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAACGA 1800
DB 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAACGA 1800
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGCGTTCACT 1860
DB 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGCGTTCACT 1860
QY 1861 TTGGCAGCTTTGCAAGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1920
DB 1861 TTGGCAGCTTTGCAAGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1920
QY 1921 GGTATTTTGTCAATTTTGTCTCAGAAATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1921 GGTATTTTGTCAATTTTGTCTCAGAAATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
QY 1981 GAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 1981 GAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 2101 CTGGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTGCTCAATTCCTTGTCTGCTCATGATG 2220
DB 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTGCTCAATTCCTTGTCTGCTCATGATG 2220

QY 2221 ACCATTGGCTTACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG 2280
DB 2221 ACCATTGGCTTACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG 2280
QY 2281 GACTGCTCTATGGTAAAAACACATTGGCTCTTTTCCACCACTGCAATCTTAAACTGC 2340
DB 2281 GACTGCTCTATGGTAAAAACACATTGGCTCTTTTCCACCACTGCAATCTTAAACTGC 2340
QY 2341 CCTGGGCTTTTCTGCT 2400
DB 2341 CCTGGGCTTTTCTGCT 2400
QY 2401 ATTAAGTTTATCT 2460
DB 2401 ATTAAGTTTATCT 2460
QY 2461 ATCTTGTTCATCT 2520
DB 2461 ATCTTGTTCATCT 2520
QY 2521 TGGACAAGATCAAAACACCCCAAGCTTGTGATGATCAATTAATCTCTGATGATGTCGAAAAACAG 2580
DB 2521 TGGACAAGATCAAAACACCCCAAGCTTGTGATGATCAATTAATCTCTGATGATGTCGAAAAACAG 2580
QY 2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 2641 CCTCCAGTTCGCTGCCATCAAGCTTATCCAGTGTGAGAGCTGCGCATCTTTCTCTCT 2700
DB 2641 CCTCCAGTTCGCTGCCATCAAGCTTATCCAGTGTGAGAGCTGCGCATCTTTCTCTCT 2700
QY 2701 GTGGCAATTTGCTCCCATGCTCTCTAA 2724
DB 2701 GTGGCAATTTGCTCCCATGCTCTCTAA 2724

Search completed: July 12, 2005, 14:17:31
Job time : 1585.87 secs

This page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2724	100.0	2724	4	US-09-170-496D-263	Sequence 263, Appl
2	2724	100.0	4570	4	US-09-376-594-201	Sequence 201, Appl
3	2720.8	99.9	2724	4	US-09-170-496D-277	Sequence 277, Appl
4	710.4	26.1	723	4	US-09-495-050A-220	Sequence 220, Appl
5	556.4	20.4	2988	4	US-09-799-451-723	Sequence 723, Appl
6	540	19.8	2612	4	US-09-495-050A-214	Sequence 214, Appl
7	511.2	18.8	4203	2	US-08-866-757-1	Sequence 1, Appli
8	511.2	18.8	4203	3	US-09-153-593-1	Sequence 1, Appli
9	125.8	4.6	2703	4	US-09-482-273-75	Sequence 75, Appl
10	123.8	4.5	473	2	US-08-866-757-3	Sequence 3, Appli
11	123.8	4.5	473	3	US-09-153-593-3	Sequence 3, Appli
12	115.4	4.2	2709	4	US-09-482-273-101	Sequence 3, Appli
13	106.8	3.9	2413	4	US-09-016-434-1459	Sequence 1459, Ap
14	100.2	3.7	3710	4	US-09-186-350A-54	Sequence 54, Appl
15	100	3.7	4417	3	US-07-741-453A-57	Sequence 57, Appl
16	98.6	3.6	2292	4	US-09-826-509-402	Sequence 402, Appl
17	98.6	3.6	2459	4	US-09-186-350A-54	Sequence 54, Appl
18	97	3.6	2292	4	US-09-826-509-394	Sequence 394, Appl
19	97	3.6	2292	4	US-09-826-509-398	Sequence 398, Appl
20	97	3.6	2292	4	US-09-826-509-406	Sequence 406, Appl
21	97	3.6	2292	4	US-09-826-509-414	Sequence 414, Appl
22	95.4	3.5	2292	4	US-09-826-509-410	Sequence 410, Appl
23	93.8	3.4	2292	4	US-09-826-509-418	Sequence 418, Appl
24	93.8	3.4	3549	4	US-09-826-509-586	Sequence 586, Appl
25	92.2	3.4	2292	4	US-09-826-509-422	Sequence 422, Appl
26	91	3.3	2179	1	US-08-487-886-1	Sequence 1, Appli
27	91	3.3	2179	2	US-08-531-070A-1	Sequence 1, Appli

Db ||||| 361 CTGAGAAATATCAGCTAAGACAGTACCCACAGAGCTCTGCGAATTTTCGAGGCCTT 420
Qy ||||| 421 CAATCCCTGCGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db ||||| 421 CAATCCCTGCGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy ||||| 481 CTGCAATTCCTTGAGGCACTGCTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db ||||| 481 CTGCAATTCCTTGAGGCACTGCTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Qy ||||| 541 GCTTTTAGAAGTTTATCGGCATTCGACGATGCAAGCCATGCTGGCCCTGGAACAAATACACCAC 600
Db ||||| 541 GCTTTTAGAAGTTTATCGGCATTCGACGATGCAAGCCATGCTGGCCCTGGAACAAATACACCAC 600
Qy ||||| 601 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db ||||| 601 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy ||||| 661 AGAATCCAATCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db ||||| 661 AGAATCCAATCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Qy ||||| 721 TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db ||||| 721 TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Qy ||||| 781 GAACTAGGATTTCTAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db ||||| 781 GAACTAGGATTTCTAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy ||||| 841 CCTCTCTTATTAACAATATCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
Db ||||| 841 CCTCTCTTATTAACAATATCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
Qy ||||| 901 TTTCAACATTTACCTTGAATGAACACTGACTCTGAAATGCTGCTCAACAATAACTGAA 960
Db ||||| 901 TTTCAACATTTACCTTGAATGAACACTGACTCTGAAATGCTGCTCAACAATAACTGAA 960
Qy ||||| 961 TTTCTCGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
Db ||||| 961 TTTCTCGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
Qy ||||| 1021 TCATCTCTCTCAAAACCGCTGCAATCAGTTACCTAATCTCAAGTCTAGATCTGCT 1080
Db ||||| 1021 TCATCTCTCTCAAAACCGCTGCAATCAGTTACCTAATCTCAAGTCTAGATCTGCT 1080
Qy ||||| 1081 TACAACCTATTAGAGATTTTACCAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
Db ||||| 1081 TACAACCTATTAGAGATTTTACCAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
Qy ||||| 1141 CTAAGACATAATGAATCTAGCAAAATTAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Db ||||| 1141 CTAAGACATAATGAATCTAGCAAAATTAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Qy ||||| 1201 CGATCGCTGAATTTGGCTTGAACAAATTTGCTATATTACCCCAATGCAATTTTCCAAT 1260
Db ||||| 1201 CGATCGCTGAATTTGGCTTGAACAAATTTGCTATATTACCCCAATGCAATTTTCCAAT 1260
Qy ||||| 1261 TTGCCATCCCTTAAAGCTGACCTATGCTCCAACTCTGCTGCTCTTTTCTCTATAACT 1320
Db ||||| 1261 TTGCCATCCCTTAAAGCTGACCTATGCTCCAACTCTGCTGCTCTTTTCTCTATAACT 1320
Qy ||||| 1321 GGGTTACATGGTTTAACTCACTTAAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db ||||| 1321 GGGTTACATGGTTTAACTCACTTAAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Qy ||||| 1381 TCATCTGAAAATTTCCAGAACTCAGGTTTATAGAAATGCTTATGCTTACAGTGGCTGT 1440
Db ||||| 1381 TCATCTGAAAATTTCCAGAACTCAGGTTTATAGAAATGCTTATGCTTACAGTGGCTGT 1440
Qy ||||| 1441 GCATTTGGAGTGTGAGAAATGCCCTTATAGAAATTTCTAATCAATGGAATTAAGGTTGACAA 1500

Db GCATTTGGAGTGTGAGAAATGCCCTTATAGAAATTTCTAATCAATGGAATTAAGGTTGACAA 1500
Qy ||||| 1501 AGCAGTATGAGACGACCTTCTATAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db ||||| 1501 AGCAGTATGAGACGACCTTCTATAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Qy ||||| 1561 GACCTTCAAGATTTCTGCTTGTGAGTGAAGACCTGAAAGCCCTTCATTTCAAGTGCAG 1620
Db ||||| 1561 GACCTTCAAGATTTCTGCTTGTGAGTGAAGACCTGAAAGCCCTTCATTTCAAGTGCAG 1620
Qy ||||| 1621 TGTTCACCTTTCCAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGGCTGGCTGATC 1680
Db ||||| 1621 TGTTCACCTTTCCAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGGCTGGCTGATC 1680
Qy ||||| 1681 AGAATGGAGTGTGGACCAATAGCAGTTCTGGCACTTACTTGTATATGCTTTGGTGACTTCA 1740
Db ||||| 1681 AGAATGGAGTGTGGACCAATAGCAGTTCTGGCACTTACTTGTATATGCTTTGGTGACTTCA 1740
Qy ||||| 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACCTGTTAAATTTGGGGTTCATCGCA 1800
Db ||||| 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACCTGTTAAATTTGGGGTTCATCGCA 1800
Qy ||||| 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGTGTGATGGCTTCACT 1860
Db ||||| 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGTGTGATGGCTTCACT 1860
Qy ||||| 1861 TTTGGCAGCTTTGACACATGCTGCTGGTGGAGAAATGGGGTTGGTGGCATGTCAAT 1920
Db ||||| 1861 TTTGGCAGCTTTGACACATGCTGCTGGTGGAGAAATGGGGTTGGTGGCATGTCAAT 1920
Qy ||||| 1921 GGTTTTGTGCTCAATTTTGTTCAGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2040
Db ||||| 1921 GGTTTTGTGCTCAATTTTGTTCAGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTCTAGC 2040
Qy ||||| 2041 CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGCGCTTGACCATGGCGGAGTTCCCTGT 2100
Db ||||| 2041 CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGCGCTTGACCATGGCGGAGTTCCCTGT 2100
Qy ||||| 2101 CTGGGTGGCAGAGTATGGCGCTTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGGCC 2160
Db ||||| 2101 CTGGGTGGCAGAGTATGGCGCTTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGGCC 2160
Qy ||||| 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTTCCTCATGATG 2220
Db ||||| 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTTCCTCATGATG 2220
Qy ||||| 2221 ACCATGGCTTACCAAGCTCTACTGCAATTTGGAAGGGAGACCTGGAGAAATTTTGG 2280
Db ||||| 2221 ACCATGGCTTACCAAGCTCTACTGCAATTTGGAAGGGAGACCTGGAGAAATTTTGG 2280
Qy ||||| 2281 GACTGCTCTATGTTAAACACATTTGCTCTTCCAACTGCTTCCAACTGCTTAAACTGC 2340
Db ||||| 2281 GACTGCTCTATGTTAAACACATTTGCTCTTCCAACTGCTTCCAACTGCTTAAACTGC 2340
Qy ||||| 2341 CCTGGGCTTTCTTGTCTCTCTCTTTTAAATAAACCTTACATTTTACGTCCTGGAAGTA 2400
Db ||||| 2341 CCTGGGCTTTCTTGTCTCTCTCTTTTAAATAAACCTTACATTTTACGTCCTGGAAGTA 2400
Qy ||||| 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTGCTGATGCTCAATCCCCTTCTCTAC 2460
Db ||||| 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTGCTGATGCTCAATCCCCTTCTCTAC 2460
Qy ||||| 2461 ATCTTGTTCATCTCACTTTAAGGAGATCTGGTGGCTTGGAGAAAGCAAACTCAAGTC 2520
Db ||||| 2461 ATCTTGTTCATCTCACTTTAAGGAGATCTGGTGGCTTGGAGAAAGCAAACTCAAGTC 2520
Qy ||||| 2521 TGGACAGATCAAAACACCCCAAGCTTGAATGCTCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db ||||| 2521 TGGACAGATCAAAACACCCCAAGCTTGAATGCTCAATTAATCTCTGATGATGTCGAAAAACAG 2580

[illegible]

D _b	1621	TGTTCACTTCCCACGGCCCTTCAAACCCTGTGAACAACCTGCTTGATNGCGTGGCTGATC	1680
Q _y	1681	AGAAATTGGAGTGTGGACCATAGCAGTTCTTG6CACTTACTTTGTAATGCTTTGGTGACTTCA	1740
D _b	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTTG6CACTTACTTTGTAATGCTTTGGTGACTTCA	1740
Q _y	1741	ACAGTTTTGCAGATCCCCCTCTGTGACATTTCCGCCCATTTAAACTGTTAATTTGGGGTCAATCGCA	1800
D _b	1741	ACAGTTTTGCAGATCCCCCTCTGTGACATTTCCGCCCATTTAAACTGTTAATTTGGGGTCAATCGCA	1800
Q _y	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGC CGTGTCTGGCTGGTGGATGCGTTTCACT	1860
D _b	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGC CGTGTCTGGCTGGTGGATGCGTTTCACT	1860
Q _y	1861	TTTGGCAGCTTTGCACGACATGGTGCCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAATT	1920
D _b	1861	TTTGGCAGCTTTGCACGACATGGTGCCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAATT	1920
Q _y	1921	G GTTTTTTGTGCATTTTTTGTCTCAGAATCAATCTGTTTTTCTGTCTTACTCTGGCAGCCCTG	1980
D _b	1921	G GTTTTTTGTGCATTTTTTGTCTCAGAATCAATCTGTTTTTCTGTCTTACTCTGGCAGCCCTG	1980
Q _y	1981	GAGCGTGGTTCCTGTGGAATAATTCGCAAAAATTTGAAACGAAAGCTCCATTTTCTTAGC	2040
D _b	1981	GAGCGTGGTTCCTGTGGAATAATTTGCAAAAATTTGAAACGAAAGCTCCATTTTCTTAGC	2040
Q _y	2041	CTGAAAGTAATCAATTTTTGTCTCTGTGCCCTCTGTGCCCTTGACCATGGCCGCAGTTCCCTCG	2100
D _b	2041	CTGAAAGTAATCAATTTTTGTCTCTGTGCCCTCTGTGCCCTTGACCATGGCCGCAGTTCCCTCG	2100
Q _y	2101	CTGGTGTGCACGAAGTATGCGGCTCCCTCTCTGTGCTGCTTTTGCTTTTGGGAGCCCC	2160
D _b	2101	CTGGTGTGCACGAAGTATGCGGCTCCCTCTCTGTGCTGCTTTTGCTTTTGGGAGCCCC	2160
Q _y	2161	AGCACCATGGGTACATGGTGBCTCTCATCTTGCTGAAATTCCTTTGCTTCCCTCATGATG	2220
D _b	2161	AGCACCATGGGTACATGGTGBCTCTCATCTTGCTGAAATTCCTTTGCTTCCCTCATGATG	2220
Q _y	2221	ACCATTGCTACACAAAGCTCTACTGCAAATTTGGACAAGGAGACCTGGAGAAATATTGG	2280
D _b	2221	ACCATTGCTACACCAAGCTCTACTGCAAATTTGGACAAGGAGACCTGGAGAAATATTGG	2280
Q _y	2281	GACTGCTCTATG GTTAAACA CATTTG CCCTGTCTTCAACCAAGCTCATCTTAAACTGC	2340
D _b	2281	GACTGCTCTATGAAAAAACATTTGCCCTGTGTCTTCTTCAACCAAGCTCATCTTAAACTGC	2340
Q _y	2341	CCTGTGGCTTCTTGCTCTTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCGAAGTA	2400
D _b	2341	CCTGTGGCTTCTTGCTCTTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCGAAGTA	2400
Q _y	2401	ATTAAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCTTCTCTAC	2460
D _b	2401	ATTAAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCTTCTCTAC	2460
Q _y	2461	ATCTTGTTCAATCCTCACTTTTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAAAACCTACGTC	2520
D _b	2461	ATCTTGTTCAATCCTCACTTTTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAAAACCTACGTC	2520
Q _y	2521	TGGACAAGATCAAAACAACCCCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAANAACAG	2580
D _b	2521	TGGACAAGATCAAAACAACCCCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAANAACAG	2580
Q _y	2581	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACACAGCTCCAGACATCACTTATGACCTG	2640
D _b	2581	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACACAGCTCCAGACATCACTTATGACCTG	2640
Q _y	2641	CCTCCCAGTTCCCGTGCCATCACCAGCTTATCCAGTGA CTGAGAGCTGCCATCTTTCTCTCT	2700
D _b	2641	CCTCCCAGTTCCCGTGCCATCACCAGCTTATCCAGTGA CTGAGAGCTGCCATCTTTCTCTCT	2700
Q _y	2701	GTGGCATTTGTGCCATGTCTTAA	2724
D _b	2701	GTGGCATTTGTGCCATGTCTTAA	2724

QY 528 AATCCCGCTCAGGCTTTTGNAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAA 587
Db 172 GGTGCTGTGCAACCCCTCAGCAATCTGCCACCCCTACAGGCGTGACCCCTGGCTCTCAA 231
QY 588 CAAATACACACATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTGGTAGTTCTACA 647
Db 232 CAAGATCTCAAGCATCCCTGACTTTTGCATTTTACCAACCTTTCAAGCCCTGGTAGTTCTGCA 291
QY 648 TCTCATTAACAATAGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTTCCAGGCT 707
Db 292 TCTTCATAAACAATAAATTTAGAGCCTGAGTCAACACTGTTTGGTGACTAGATAACCT 351
QY 708 AGAGACTTTAGATTTAAATTAACATAAATCTTGATGAATTTCCCACTGCAATTAAGCACT 767
Db 352 GGAGACCTTAGACTTTGAATTTAAATTTTAACTTTGGGGGAATTTCCCTCAGGCTATTAAGGCCCT 411
QY 768 CTCGAACCTTAAAGAACTTAGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGC 827
Db 412 TCCTAGCCCTTAAGAGCTAGATTTTATAGTAATTTCTGTTATCCCTGATGGAGC 471
QY 828 ATTTGTAGGCAACCTTCTCTTATTAACAATACATTTTCTATGACAAATCCCAATTTGT 887
Db 472 ATTTGATGGTAATCCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTGT 531
QY 888 TGGGAGATCTGTTTCAACATTTTACCTGAATTAAGAACACTGACTCTGAATGGTGCCTC 947
Db 532 GGGGAACCTCAGCATTTTCAAAATTTTATCTGATCTTCAATTTCCCTAGTCAATCTGGTGGCAAG 591
QY 948 ACAATAAATCTGAATTTTCCCTGATTTAACTGGAACCTGCAACCTGGAGAGCTGACTTTAAAC 1007
Db 592 CATGGTGAGCAGTTTCCCAATCTTACAGGAACCTGTCACCTGGAAAGCTCTGACTTTGAC 651
QY 1008 TGGAGCAGATCTCATCTCTCTCTTCAACCGTCTGCAATCAGTTTACCTAATCTCCAAGT 1067
Db 652 AGGTACAAGATTAAGCAGCATACCTAAATAATTTGTGCCAAGAACAAAGATGCTTAGGAC 711
QY 1068 GCTAGATCTGCTTACACCTTATAGAGATTTTACCGAGTTTTCAGTCTGCCAAGCT 1127
Db 712 TTTGGACTTGTCTTACAAATAATTAAGAGACCTTCCAAAGTTTAAATGGTTGCCATGCTCT 771
QY 1128 TCAGAAAATTCGACCTTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCA 1187
Db 772 GGAAGAAATTTCTTACAGCGTAACTCAATCTACCAATTAAGGAAGCACCCTTCAAG 831
QY 1188 GTTGTAGCTTCGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTTATTCACCCCAA 1247
Db 832 CCTGATATCTCTAAGGATTTAGATCTGAGTAGAAACCTGATACATGAATAATTCACAGTAG 891
QY 1248 TGCAATTTTCCACTTTGGCATCCCTAATAAGCTGGACCTATCGTCCCACTCTCTGCTC 1307
Db 892 AGCTTTTCCACACTTTGGCCCAATTAACCTAGATGAATTTCAATGAATTAACCTTC 951
QY 1308 TTTTCTATAACTGGGTTACATGGTTTAACTCACTTTAAATTAACAGGAAATCATGCTT 1367
Db 952 CTTTCTACGAGGCGCTGAATGGGCTAAATCACTGAATCTGTGGGCACTTCAAGCT 1011
QY 1368 ACAGAGCTTGATATCATCTGAATACTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGC 1427
Db 1012 GAAAGAGCCTTAGCAGCAAAAAGACTTTGTTAACTCAGGCTTTTATCAGTACCATATGC 1071
QY 1428 TTACCGAGCTGTGCATTTGGAGTGTGTGAGAGTCCCTATAAGATTTTCTAATCAATGAA 1487
Db 1072 TTATCAGTGTGTGATTTTGGGGTTGTG-----ACTCTTATGCAAAATTTAA 1119
QY 1488 TAAAGTGACAAACAGCAGTATGGAGCAGCTTTCATAAGAAAGATGCTGGAATGTTTCAGGC 1547
Db 1120 CACAGAAATTAACAGCTTCCAGGACCAC-----AGTGTGGCAGAGA 1161
QY 1548 TCAAGATGAACGTGACCTTGAAGATTTCTGCTTGAATTTTGGAGGAAGACCTGGAAGCCCT 1607
Db 1162 GAAAGGTACTGTGATGAGCAAAATGTCAAGACACTCTTGAAAAATGAAGACATAGTCA 1221

QY 1608 TCATTCAGTGCAGTGTTCACCTTCCCGAGCCCTTCAAAACCTGTGTGAACACCTGCTTGA 1667
Db 1222 AATAATTTATCCATTTGTACACCTTCAACAGGTGCTTTTAAAGCCCTGTGNAATTTACTGGG 1281
QY 1668 TGGCTGGCTGATCAGAAATTTGGAGTGTGGAACATAGCAGTCTTGGACACTTACTGTAAATGC 1727
Db 1282 AAGCTGGATGATTCGTCTACTGTGTGTTCAATTTCTTGGTGTCAATTTATTTTCAACCT 1341
QY 1728 TTTGGTGACTTCAACAGTTTTTCAGATCCCTCTCTGTAATTTTCCCACTTAACTGTTAAT 1787
Db 1342 GCTTGTATTTTAAACAACATTTGCACTCTGATCATCTGCTTTCGTCCAAATTTGTTAT 1401
QY 1788 TGGGCTCATCCAGCAGTGAACATGCTCAGGGAGTCTCAGATGCCGTGCTGGCTGGTGT 1847
Db 1402 AGGCTTGATTTCTGTGCTAACTTATTCATGGGAATCTATACTGGCATCTCTAACTTTCT 1461
QY 1848 GGATGCGTTCACTTTTGGCAGCTTTTGACGCAATGTTGTCCTGTTGGGAGAAATGGGTTGG 1907
Db 1462 TGATGCTGTGCTTGGGGCAGATTCGCTGAATTTGGCAATTTGGTGGGAAACCTGCGAGTGG 1521
QY 1908 TTGCCATGTCAATGGTGTTCATTTTGTTCATTTTGTCTCAGAACTCATCTGTTTCTGCTTAC 1967
Db 1522 CTGCAAAAGTACCTGGGTTTCTTGCAGTTTCTCTCAGAAAGTGCATATTTTATTAAT 1581
QY 1968 TCTGGCAGCCCTGAGCGTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAAACGAAAGC 2027
Db 1582 GCTAGCAACTCTGAAAGAGCTTATCTGCAAAAGATATATGAAATTTGGAAGACAA 1641
QY 2028 TCCATTTTCTAGCCTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTTGAACAATGGC 2087
Db 1642 TCATCTCAACAGTTTCGGGTGTGCTGCTTTCCTAGGTGCTACAGTAGCAGG 1701
QY 2088 CGCAGTTTCCCTGCTGGGTGGCAGCAAGTATGGGCTCTCCCTCTCTGCTGCTTGGCC 2147
Db 1702 CTGTTTTTCCCTTTTCCATAGAGGGGAATATTCGCAATCACCCCTTTGTTGGCAATTTCC 1761
QY 2148 TTTTGGGAGCCGAGCACCAGCTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTTG 2207
Db 1762 TACAGGTGAAAGCCATCATTAGGATTCACCTGTAAGCTTAGTGTATTTAACTCACTAGC 1821
QY 2208 CTTCTCATGATGACCAATTTGCTTACACCAAGCTCTACTGCAATTTTGGCAAGGAGACCT 2267
Db 1822 ATTTTATTAATGGCCGTATCTACATAAGCTATACTGCAACTTGGCAAAAGAGGACCT 1881
QY 2268 GGAGAAATTTGGGACTGCTCTATGTTAAACACATTTGCCCTGTGCTTCCCAACTG 2327
Db 1882 CTCAGAAACTCAACAATCTAGCATGATTAAGCATGTGCTTGGCTTAATCTTCAACAATG 1941
QY 2328 CATCTAAACTGCTGCTGCTTTCTTGTCTCTCTCTCTCTCTTTTAAATAAACCTTACATTTAT 2387
Db 1942 CATCTTTTCTGCTGCTGCTGCTGCTTTTTCATTTGCACTTATGATCACTGCAATCTCTAT 2001
QY 2388 CAGTCTGAAAGTAATTAAGTTTATCTCTGCTGGTGGTAGTCCCACTTCTGCTGATGTCTCAA 2447
Db 2002 CAGCCCGAAATAATGAATGCTGTACTCTGATATTTTTCATTTTTCATTTGCTGCTGCTGCA 2061
QY 2448 TCCCTCTCTCTACATCTTGTTCATCTCTTAAAGAGATCTGTTGAGCCTGA 2503
Db 2062 TCCAGTCTGTATGTTTCTTCAACCCCAAGTTTAAAGAGACTGGAAGTTACTGA 2117

RESULT 7

US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA

STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4203 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-866-757-1

Query Match 18.8%; Score 511.2; DB 2; Length 4203;

Best Local Similarity 54.9%; Pred. No. 2.2e-146; Mismatches 863; Indels 31; Gaps 3;

Qy	520	TTAACAAGAAATCCCGTTCAGGCTTTTAAAGATTTATGGGATTCGACATTCGAAAGCCATGACCTTG	579
Db	56	TTGACGGAGGTGCTGTGCACCCCTCAGCAATCTGCCACCCATACAGGCGCTGACCCCTG	115
Qy	580	GCCCTGAACAAATACACACATACAGACTATGCTTTGGAACTCTCCAGCTTGCTA	639
Db	116	GCTCTCAACAAGATCTCAAGCATCCCTGACTTTGCAATTTACCACTTTCAAGCCCTGTA	175
Qy	640	GTTCTACATCTCCATAACAATAGATCCATCCCTGGGAAGAAATGCTTTGATGGGCTC	699
Db	176	GTTCTGCATCTTCATAACAATAAATTAAGAGCTGAGTCAACACTGTTTGTATGGACTA	235
Qy	700	CACAGCTAGAGACTTTAGATTTAAATTAACAATACCTTTGATGAATTCGCCACTGCAAT	759
Db	236	GATACTTGGAGACCTTAGACTTGAATTAATAACTTTGGGGAAATTTCTCAGGCTATT	295
Qy	760	AGGACACTCTCAACCTTAAGAACTAGAGTTTCATAGCAACAATATCAGGTCGATCCT	819
Db	296	AAAGCCCTTCTAGGCTTTAAGAGCTAGAGTTTCATAGTAAATCTTCTTCTGTTATCCCT	355
Qy	820	GAGAAGCAATTTGTAGGCAACCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATC	879
Db	356	-ATGAGAGCAATTTGATGGTAATCCACTTTAAGAACTATACATTTGTATGATTAATCCTCTG	414
Qy	880	CAATTTGTTGGAGATCTGCTTTTCAACATTTACCTGAACTAAGAACTGACTCTGAAT	939
Db	415	TCCTTTTGGGAACTCAGCAATTCACAAATTTATCTGATCTTCATCTCCCTAGTCAATCGT	474
Qy	940	GCTGCTCACAATAACTGAATTTCTGATTTAACTTGAACCTGCAAACTCTGAGAGTCTG	999
Db	475	GCTGCAAGCATGGTCAGCAGTTCCCAATCTTACAGGAATGTCCCACTTGGAAAGTCTG	534
Qy	1000	ACTTTAACTGGAGCAGATCTCTCTTCTCTCAACGGTCTGCAATCAGTTACCTAAT	1059
Db	535	ACTTTGACAGGTACAAAGATAGCAGCATACCTAATAATTTTGTGTCAAGAAACAAAGATG	594

Qy	1060	CTCCAAGTGTAGATCTCTCTTACAACTATTAGAAATTTACCAGTTTTTTCAGTCTGC	1119
Db	595	CTTAGGACTTTGGACTTGTCTTACAAATAATATAAGAGACCTTCCAAATTTTAAATGTTGC	654
Qy	1120	CAAAAGCTTCAGAAATTCAGTAAGACATTAATGAATCTACGAATTAAGTTGACACT	1179
Db	655	CATGCTCTGGAAGAAATTTCTTTACAGCGTAATCAAAATCTACCAATAAAGGAAGCACC	714
Qy	1180	TTCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTAT	1239
Db	715	TTTCAAGCCCTGATATCTCTAAGGATTTAGATCTGAGTAGAAGAACTGATACATGAAT	774
Qy	1240	CACCCCAATGATTTTCCACTTTCATCCCTTAATAAGCTGAGACCTATGTCACACCTC	1299
Db	775	CACAGTAGAGCTTTTGGCCACACTTGGGCCAATAACTAACTAGATGTAAGTTTCAATGAA	834
Qy	1300	CTGTGCTCTTTTCTTATTAACCTGGGTATACATGGTTTAACTCACTTAAATTAACAGAAAT	1359
Db	835	TTAACTTCTCTTCTTACGGAAGGCTGAATGGGCTTAAATCAACTGAAACTTGGTGGCAAC	894
Qy	1360	CATGCTTTACAGAGCTTGATATCATCTCAAAACTTTCCAGAACTCAAGGTTATAGAAATG	1419
Db	895	TTCAAGCTGAAGAGCCCTTAGCAGCAAGACCTTTGTTAACCTCAGSTCTTTATCAGTA	954
Qy	1420	CCTTATGCTTACAGTGTCTGTGCATTTGGAGTGTGTGAGATGCTTAAGATTTCTAAT	1479
Db	955	CCATATGCTTATCAGTGTCTGTGCATTTTGGGGTGTG-----ACTTTATGCA	1002
Qy	1480	CAATGGAATAAGGTGACACAGCAGTATGACGACCTTCAAGAAAGATGCTGGAATG	1539
Db	1003	AATTTAAACACAGAAAGATAACAGGCTCCAGGACCACAGT-----GTG	1044
Qy	1540	TTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGTACTTTGAGGAAGACCTG	1599
Db	1045	GCACAGGAAGAGTACTGCTGATGCGCAATGTCAAGCACTCTTGAATAATGAGAA	1104
Qy	1600	AAAGCCCTTCATTCAAGTGCAGTGTTCACCTTTCCAGGAGCCCTTCAAAACCTGTGAACAC	1659
Db	1105	CATAGTCAAAATAATATCCATTGTACACTTCAACAGGTGCTTTTAAAGCCCTGTGAATAT	1164
Qy	1660	CTGCTTGTATGCTGGTGCATCAGAAATTTGGAGTGTGACCATAGCAGTCTTGGCATTACT	1719
Db	1165	TTACTGGGAAGCTGGATGATTCGCTTACTGTGTGTTTCTTTCTTGGTTGCAATATTT	1224
Qy	1720	TGTAATGCTTTGGTGACTTCAACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAA	1779
Db	1225	TTCAACCTGCTTGTATTTTAAACAATTTGCACTTGTACATCACTGCCCTTGTGTCAAA	1284
Qy	1780	CTGTTAATTTGGGGTCACTCGCAGCAGTGAAATGCTCACGGGAGTCTCCAGTGCCTGCTG	1839
Db	1285	TTGTTTATAGGCTTGAATTTCTGTGTCTAACTTATTCATGGGAATCTATCTGGCATCTTA	1344
Qy	1840	GCTGGTGGATGGTTCACCTTTTGGCAGCTTTTGCACACATGGTGCCTGTGGGAGAAAT	1899
Db	1345	ACTTTTCTTGTATGCTGTCTCTGGGCGCAGATTCGCTGAATTTGGCAATTTGGTGGGAACT	1404
Qy	1900	GGGGTTGGTGCCTATGCTATTTTGTGTCATTTTGTGCTTCAAGATCATCTGTTTTTC	1959
Db	1405	GGCAGTGGCTCAAGTAATCTGGGTTTCTTGCAGTTTTTCTCTCAGAAAGTGCCTAATTT	1464
Qy	1960	CTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAATAATTTCTGCAAAATTTGAA	2019
Db	1465	TTAATAATGCTAGCAACTGTGCAAGAAAGCTTATCTGCAAAAGATATAATGAATAATGG	1524
Qy	2020	ACGAAGCTCCATTTTCTAGCCTGAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTG	2079
Db	1525	AAGAGCAATCATCTCAAAACAGTTCGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1584
Qy	2080	ACCATGCGCGAGTTCCCTCTCTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTG	2139
Db	1585	GTAACAGGCTGTTTTCCTCCCTTTTTCATAGAGGGGAATATTCGTACATCACCCCTTTGTTG	1644


```
Qy 2140 CTTTGGCTTTTGGGAGCCAGACCATGGCTACATGGTCTCTCATCTTCTGCTCAAT 2199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1645 CCAITTCCTACAGGTGAACACGCCATCAATTAGGATTCAGTGAACGTGATGCTATTAAC 1704
Qy 2200 TCCCTTTGCTTCTCATGATGACCAATTTGCCCTACACCAAGCTCTACTGCAATTTGGACAAG 2259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1705 TCACTAGCAATTTTATTAATGGCCGTTATCTACACTAAGCTATACTGCAACTTGGAAAAA 1764
Qy 2260 GGAGACCTGGAGAAATATTTGGGACTGCTCTATGTGTAAACACATTTGCCCTGCTCTTC 2319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1765 GAGGACCTCTCAGAAAACTCAAACTAGCAATTAAGCAATGCTGCTTGGCTTAATCTTC 1824
Qy 2320 ACCAACTGCATCTAAACTGCCCTGTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCT 2379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1825 ACCAAATGCATCTTTTCTGCCCCGTGGCGTTTTTTCATTTGGCACCATTGATCACTGCA 1884
Qy 2380 ACATTTATCAGTCTCTGAAGTAATTAAGTTTATCTCTCTGCTGGTAGTCCCACTTCTCTGCA 2439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1885 ATCTATACGCCCGGAAATTAAGTCTGTTACTCTGATATTTTTCATTTCCATGCTGCT 1944
Qy 2440 TGTCTCAATCCCTTCTCTCATCTTTGTTCAATCTCTCACTTTAAGGAGATCTGGTGAGC 2499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1945 TGCCTGAATCCAGTCTGTATGTTTCTTCAACCAAAAGTTTAAAGAAGACTGGAAGTTA 2004
Qy 2500 CTGA 2503
Db |||||
2005 CTGA 2008

RESULT 8
US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGMAN, DEK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4203
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-153-593-1

Query Match 18.8%; Score 511.2; DB 3; Length 4203;
Best Local Similarity 54.9%; Pred. No. 2.2e-146;
Matches 1090; Conservative 0; Mismatches 863; Indels 31; Gaps 3;

Qy 520 TTAACAGAAATCCCGTCCAGGCTTTTAGAGTTTATGGCAATTCGCAAGCCATGACCTTG 579
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 TTGACGGAGGTGCGTGTGCACCCCTCAGCAATCTGCCACCCCTACAGGCGCTGACCCGTG 115
Qy 580 GCCCTGAACAAAATACACCATACAGACTATGCTTTTGGAAACCTCTCCAGCTTGGTA 639
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 GCTCTCAACAAGATCTCAAGCATCCGCTGACTTTGGATTTACCAACCTTCAAGCTGGTA 175
Qy 640 GTTCTACATCTCCATAACAATAGAAATCCATCCCTGGGAAAGAAATGTTTGTATGGGCTC 699
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 GTTCTGCACTTCTAACAATAAATAAAGAGAGCTGAGTCAACACTGTTTGTATGGACTA 235
Qy 700 CAGACCTTAGACACTTTAGATTTAAATTAATTAACCTTGATGATTTCCCACTGCAAT 759
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 GATAACCTGGAGACCTTAGACTTGAATTTAATAAATTTGGGGAAATTTCCCTCAGGCTATT 295
Qy 760 AGGACACTCTCAACCTTAAAGAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCT 819
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 296 AAAGCCCTTCTAGCCCTTAAAGAGCTAGGATTTTCATAGTAATTCATTTCTGTGTTATCCCT 355
Qy 820 GAGAAACATTTGTAGCAACCCCTTCTCTTATTACAAATACATTTCTATGACAAATCCCATC 879
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 -ATGAGACATTTGATGTTAATCCACTCTTAAAGAACTATACATTTGTATGATTAATCCCTCG 414
Qy 880 CAATTTGTTGGAGATCTGCTTTTCAACATTTTACCTGAACTTAAGAAACACAGTCTGAAAT 939
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 TCTTTGTGGGAACTCAGCAATTTCAAAATTTATCTGATCTTCAATCCCTAGTCAATTCGT 474
Qy 940 GGTGCTCTACAAATAACTGAATTTCTGATTTAACTGAACTGCAAACTGGAGAGTCTG 999
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 GGTGCAAGCATGTTGAGCAGTTCCTCCAACTTACAGGAACTGTCACCTGGAAAGTCTG 534
Qy 1000 ACTTTAACTGGAGACAGATCTCATCTCTCTCTCAAAACCGTCTGCAATCAGTTACCTAAT 1059
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
535 ACTTTGACAGTACAAGATAAGCAGCATACCTAATAATTTGTGTCAAGAACAAAAAGATG 594
Qy 1060 CTCCAAGTGTAGATCTGTCTTACCACTTATTAGAAGATTTACCAGTCTTTTTCAGTCTGC 1119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
595 CTTAGGACTTTGGACTTGTCTTACAAATAATAAGAGACCTTCCAAGTTTAAATGGTTCG 654
Qy 1120 CAAAAGCTTCAGAAAAATGACCTAAGACATAAATGAAATCTACGAAATTTAAAGTTGACACT 1179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
655 CATGCTCTGGAAGAAATTTCTTTACAGCGTAATCAAACTTACCAATAAAGAAAGGCACC 714
Qy 1180 TCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGCTGTTGGAAACAAATTTGCTATTATT 1239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
715 TTTCAAGGCCCTGATATCTCTAAGGATTTAGATCTGAGTAGAAACCTGNTACATGAATTT 774
Qy 1240 CACCCCAATGCATTTTCCACTTTGCCATCTCTAATAAAGCTGGACCTATCGTCCAACTC 1299
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
775 CACAGTAGACTTTTGGCACACTTGGGCCAATACTAACTAGATGTAGTTTCAATGAA 834
Qy 1300 CTGTGCTTTTCTATATACTGGTTACATGTTTAACTCACTTAAATTAACAGGAAAT 1359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
835 TTAACCTTCTCTCTCGGAAGGCTGAAATGGGCTAAATCAACTGAACTGGTGGGCAAC 894
Qy 1360 CATGCTTACAGACTTGATATCATCTGAAACCTTTCCAGAACTCAAGTTTATAGAATG 1419
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
895 TTCAAGCTGAAGAAGCCTTTAGCAGCAAAAGACTTTGTAACTCAGGTCTTTATCAGTA 954
Qy 1420 CTTTATGCTTACCAAGTCTGTGCAATTTGGAGTGTGTGAGAATGCTATAAGATTTCTAAT 1479
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
955 CCATATGCTTATCAGTGTCTGCAATTTGGGGTTG-----ACTCTATGCA 1002
Qy 1480 CAATGGAATAAGGTGACAAACAGCAGTATGGACACTTCATAAGAAAGATGCTGGAAATG 1539
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1003 AATTTAAACACAGAAAGATAACAGCCTCCAGACCACACAGT-----GTG 1044
Qy 1540 TTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTG 1599
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1045 GCACAGGAAAGAGTACTGCTGATGCAAGCAAAATGCTCAAGCACTCTTGAATAAGAA 1104
Qy 1600 AAAGCCCTTCACTCAGTGCAGTGTTCACCTTCCCAGGCCCTTCAAAACCTCTGAAACAC 1659
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1105 CATAGTCAATAATATTCATTGTACCTTCAACAGTGTCTTAAAGCCCTGTGAATAT 1164
Qy 1660 CTGCTTGATGCTGGCTGATCAGAAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACT 1719
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1165 TTAAGGAAAGCTGGATGATTCGTCTTACTGTGTGGTTCAATTTTCTGGTTGCAATTTT 1224
Qy 1720 TGTAACTGTTTGGTGAATCTCAAGATTTTCAGATCCCTCTGTATCAATTTCCCATTA 1779
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1225 TTCAACCTGCTGTTTATTTTAAACAAATTTGCACTCTGTATCATCACTGCTTCCTCA 1284
Qy 1780 CTGTTTAAATTTGGGTCATCGCAGCAGTGAACATGCTCAGGGAGTCTCAGTGCCTGCTG 1839
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1285 TTGTTTATAGCTTGATTTCTGTGTCTAATTTATGGAATCTATATCTGATCTCTA 1344
Qy 1840 GCTGTTGGAGTGGCTTCACTTTTGGAGCTTTTCCACAGACATGCTGCTGGTGGAGAA 1899
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1345 ACTTTTCTGATGCTGTCTCTGGGGCAGATTCGCTGAATTTTGGCAATTTGGTGGGAACT 1404
```



```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-866-757-3

Query Match          4.5%; Score 123.8; DB 2; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACGTAAACGTTAGTGTATTAACTCACTAGCATTTTATTAAATGCGCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAGAGAGACCTCTCGAATAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGCTTTCACCAACTGCATCTAACTGCCCTGTGG 2347
Db 140 GCATGATTAAGCATGCTCGCTGGCTAACTCTTCCCAATTTGCAATTTGCTTTCGCCCTGTGG 199

Qy 2348 CTTTCTTCTCTCTCTCTCTCTTAAATAAACCTTACATTTTATCAGTCTCTGAAATTAAT 2407
Db 200 CGTTTTTTTTCATTTGACCAATTTGATCACTGCAATCTCTATCAGCCCGGAAATAATGAAGT 259

Qy 2408 TTAATCCTTCTGGTGTAGTCCACTTCTCTGCATGCTTCAATCCCTTCTCTACATCTTGT 2467
Db 260 CTGTACTCTGATATTTTTTCCATTTGCTGCTTCCCTGATCCAGTCTCTGATGTTTTCT 319

Qy 2468 TCAATCTCTCACTTTAAGGAGGATCTGG 2494
Db 320 TCAACCCAAAGTTTAAAGAGGACTGG 346

RESULT 12
US-09-482-273-101
; Sequence 101, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-101

Query Match          4.2%; Score 115.4; DB 4; Length 2709;
Best Local Similarity 59.3%; Pred. No. 3.4e-24;
Matches 214; Conservative 0; Mismatches 146; Indels 1; Gaps 1;

Qy 2143 TTGCTTTTGGGGAGCCAGCACCATGAGGCTTACATGGCTCTCTCATGATGATGCTCAATTC 2202
Db 10 TTTCCTACAGGTGAAAGCCCATCATTAGGATTCATCTGTAACGTTAGTCTATTAACTCA 69

Qy 2203 CTTTGTCTCTCATGATGACCATTTGCCCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGA 2262
Db 70 CTAGCATTTTTTAAATGGCGGTTATCTACACTAAGCTATATCTGCAACTTGGAAAAAGAG 129

Qy 2263 GACCTGGAGATAATTTGGGACTGCTCTATGTAAAAACACATTTGCCCTTGTCTTCTCACC 2322
Db 130 GACCTCTCAGAAACTCAATCTAGCATGATTAAGCATGTCGCTTGGCTAATCTTACC 189

Qy 2323 AACTGCATCTTAAACTGCGCTTCTTGTGCTTCTCTCTCTTTTAAATTAACCTTACA 2382
Db 190 AATTGCATCTTTTCTGCCCTGTGGCGTTTTTTTCAITTTGCACCATTTGACCTGCAATC 249

Qy 2383 TTTATCAGTCTCTGAAGTAATTAAGTTTATCTCTTGGTGGTAGTCCCACTTCTCTGCAATG 2442
Db 250 TCTATCAGCCCGGAAATAAAGTCTGTACTCTGATATTTTTTCCA-TGCTGCTTGC 308

Qy 2443 CTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTAAAGGAGGATCTGGTGAAGCTG 2502
Db 309 CTGAATCAGTCTCTGTATGTTTCTTCAACCCAAAGTTTAAAGAGACTGGAAGTTACTG 368

Qy 2503 A 2503
Db 369 A 369

Query Match          4.5%; Score 123.8; DB 3; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACGTAAACGTTAGTGTATTAACTCACTAGCATTTTATTAAATGCGCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAGAGAGACCTCTCGAATAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGCTTTCACCAACTGCATCTAACTGCCCTGTGG 2347
Db 140 GCATGATTAAGCATGCTCGCTGGCTAACTCTTCCCAATTTGCAATTTGCTTTCGCCCTGTGG 199

Qy 2348 CTTTCTTCTCTCTCTCTCTTAAATAAACCTTACATTTTATCAGTCTCTGAAATTAAT 2407
Db 200 CGTTTTTTTTCATTTGACCAATTTGATCACTGCAATCTCTATCAGCCCGGAAATAATGAAGT 259

Qy 2408 TTAATCCTTCTGGTGTAGTCCACTTCTCTGCATGCTTCAATCCCTTCTCTACATCTTGT 2467
Db 260 CTGTACTCTGATATTTTTTCCATTTGCTGCTTCCCTGATCCAGTCTCTGATGTTTTCT 319

Qy 2468 TCAATCTCTCACTTTAAGGAGGATCTGG 2494
Db 320 TCAACCCAAAGTTTAAAGAGGACTGG 346

RESULT 11
US-09-153-593-3
; Sequence 3, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGMA, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 473
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (397) (400) (432) (459)
US-09-153-593-3

Query Match          4.5%; Score 123.8; DB 3; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACGTAAACGTTAGTGTATTAACTCACTAGCATTTTATTAAATGCGCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAGAGAGACCTCTCGAATAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGCTTTCACCAACTGCATCTAACTGCCCTGTGG 2347
```

RESULT 13

US-09-016-434-1469
 ; Sequence 1469, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1469:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2413 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g903759
 ; US-09-016-434-1469

Query Match 3.9%; Score 106.8; DB 4; Length 2413;
 Best Local Similarity 45.0%; Pred. No. 1.4e-21;
 Matches 399; Conservative 0; Mismatches 487; Indels 0; Gaps 0;

QY	1616	TCACGTGTTACCTCCCGAGCCCTTCAAAACCTGTGAACACCTGTTGATGGCTGGC	1675
DB	1257	TGGTGTGATCCCGAAGTCCGATGAGTTCAACCCGTTGGAACATAAATGGGCTACAAGT	1316
QY	1676	TCATCAGAAATGGAGTGGGACATAGCAGTTCGSCACTTACTTGTAAATGCTTTGGTGA	1735
DB	1317	TCTGAGAAATGGTGTGGTTCGTTGATGCTGCTCTCTGGGCAATGCTTTTGCC	1376
QY	1736	CTTCAACAGTTTTTCAGATCCCTCTGTACATTTCCCGCAATTAACCTGTTAATGGGGTCA	1795
DB	1377	TGCTTTATCTCTCACCAGCCACTACAACTGAACGTCCTCCCGCTTTCTCATGTGCAACC	1436
QY	1796	TGCGACAGTGAACATGCTCAGGGAGTCTCAGTGCCGCTGCTGGGTGTGGATGGCT	1855
DB	1437	TGGCCCTTTGCGAATTTCTGCATGGGATGTACCTGCTCTCATCGCCTCTGTAGACCTCT	1496
QY	1856	TCACCTTTGGCAGCTTTCACACATGGTGGTGGGAGATGGGGTGGTGGCTTGGCTG	1915
DB	1497	ACACTCACTCTGAGTACTACAAACCATGCCATCGATGGGAGACAGGCCCTGGGGTGAACA	1556

QY	1916	TCATTGGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTGCTTACTCTGGCAG	1975
DB	1557	CGGCTGGTTTTCTTCTTCTTCTTCACTGTCTTCAAGCGAGTTATCGGTGTATACGCTGACGGTCATCA	1616
QY	1976	CCCTGGAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTT	2035
DB	1617	CCCTGGAGCGCTGGTATGCGCATCACCTTGGCCATGGCCCTGACCGGAAGATCCGCTCA	1676
QY	2036	CTAGCCTGAAAGTAAATCAATTTTGTCTCTGTGCCCTCTGCTGCTGACCATGCGCGGAGTTTC	2095
DB	1677	GGCAGCATGTGCCATCATGGTTGGGGGCTGGGTGTGCTTCTCTCGCCCTGCTTTC	1736
QY	2096	CCCTGCTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTGCTTGGCTTTTGGGG	2155
DB	1737	CTTTGGTGGGAATAAGTAGTATGCCAAAGTCAGTATCTGCTGCTGCCCATGSAACCGAGA	1796
QY	2156	AGCCCGAGCACCATGGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCTCA	2215
DB	1797	CCCCCTTGTCTGGCATATATTGTTTTTGTCTGACGCTCAACATAGTTGCTTGGTCA	1856
QY	2216	TGATGACCAATTCCTTACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGGAATA	2275
DB	1857	TCGCTGCTGCTGTTTATGTGAAGATCTACATCACAGTCGGAATCCGCAACCCAGGGG	1916
QY	2276	TTTGGGACTGCTATGTGTAATAACACATGCGCTGCTGCTTCTTCCAACTGCATCCTTAA	2335
DB	1917	ACAAAGATACCAAAATTCGCAAGGATGGCTGTGTGATCTTCCCGACTTCCAGTGCA	1976
QY	2336	ACTGCCCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCCCTG	2395
DB	1977	TGGCCCCAATCTCATTTCTATGCTGTGAGCAATTTCTGAACAGCTCTCATCTGTTA	2036
QY	2396	AAGTAAATTAAGTTTATCTTCTGCTGGTGGTAGTCCCACTTCTGTCATGTCTCAATCCCTTC	2455
DB	2037	GCAACTCCAAATCTTGTGCTGCTTCTTCTTATCCAAATTAACCTCTGTGCAATCCATCC	2096
QY	2456	TCTATCATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTAGCCT	2501
DB	2097	TCTATGCTATTTCACCAAGGCTTCCAGAGGGATGTGTTTCATCCT	2142

RESULT 14

US-07-741-453A-62
 ; Sequence 62, Application US/07741453A
 ; Patent No. 6228597
 ; GENERAL INFORMATION:
 ; APPLICANT: PARMENTIER, MARC
 ; APPLICANT: LIBERT, FREDERIC
 ; APPLICANT: DUMONT, JACQUES
 ; APPLICANT: VASSART, GILBERT
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
 ; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
 ; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ; STREET: 1615 L STREET, N.W.
 ; CITY: WASHINGTON, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/741,453A
 ; FILING DATE: 19911015
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KOKULIS, PAUL N.
 ; REGISTRATION NUMBER: 16773
 ; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 3710 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-741-453A-62

Query Match 3.7%; Score 100.2; DB 3; Length 3710;

Best Local Similarity 45.3%; Pred. No. 2.2e-19;
Matches 403; Conservative 0; Mismatches 483; Indels 3; Gaps 1;

```
QY 1616 TGCAGTGTTCACCTTCCCGAGGCCCTTCAAACCTGTGAACACCTGCTGTGATGGCTGGC 1675
Db 1519 TGGTGTGTACCCCHAGTCCGATGAGTTCAACCGGTGAACACATATATGGCTACAAGT 1578
QY 1676 TGATCAGAAATGGAGTGTGGACCAATAGCAGTTCGTGGCACTTACTTGTAAATGCTTTGGTGA 1735
Db 1579 TCCTGAGAAATGGTGTGGTGTGGTGTAGTCTGTGGCTCTCTGGGCAATGCTTTGTGTC 1638
QY 1736 CTTCAACAGTTTTAGATCCCTCTGTATACATTTCCCAATTAAGTGTAAATGGGTGCA 1795
Db 1639 TGGTATTTCTCTCCACGACCACTACAAACTGAACGTCCCGCTTTCTCATGTGCAACC 1698
QY 1796 TCGCAGCAGTGAACATGCTCAGCGAGTCTCAGTGGCGTGTGGCTGGTGTGATGGT 1855
Db 1699 TGGCTTTGGGATTTCTGCATGGGATGTACTGCTCTCATCGCTCTGTAGACCTCT 1758
QY 1856 TCACTTTTGGCAGCTTTGCAGCAATGTTGCTGGTGGGAGAAATGGGTGTGGTTGCCATG 1915
Db 1759 ACACCTCACTCTGAGTACTACAACCATGCCATGCTGCGACAGACGGCCCTGGGTGCAACA 1818
QY 1916 TCATTTGGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAG 1975
Db 1819 CGGCTGGTTTTCTTCACTGCTTTGCAAGCGAGTTATCGGTGTATACGCTGACGGTCAATCA 1878
QY 1976 CCCTGGAGCTGGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTT 2035
Db 1879 CCCTGGAGCGTGGTATGCAATCACTCTGCGATGCGGCTGGACCGGAAGATCGCCTCA 1938
QY 2036 CTAGCCTGAAGATATATCTTTGCTCTGTGCGCTCTGCGCTTGACCATGCGCGAGTTC 2095
Db 1939 GGCACGATGTGCCATCATGTTGGGGGCTGGGTTGTGCTTCTCTCGCCCTGCTTC 1998
QY 2096 CCCTGCTGGTGGGAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGGCTTTTGGGG 2155
Db 1999 CTTTGGTGGGAATAGTAGTACTATGCCAAAGTCAGTATCTGCTGCCATGACACCGAGA 2058
QY 2156 AGCCAGCAGCATGGCTTACATGTCGTCTCATCTGCTCAATTTCCCTTTGCTTCCCTCA 2215
Db 2059 CCCCTCTGCTGCAATATATGTTTGTGTCAGCCTCAAATAGTGTGCTTCTGTCAT 2118
QY 2216 TGATGACCAATGCTCTACCAAGCTCTACTGCAATTTGGACAA---GGGAGACCTGGAGA 2272
Db 2119 TCGTCTGCTGCTTATGTAAGATCTATCATCACAGTCCGAAATCCGCGAGTACAACCCAG 2178
QY 2273 ATATTTGGGACTGTCTATGTTGAACACATGTCGCTTGTGCTTCTTACCAACTGCATCC 2332
Db 2179 GGGACAAAGATACCAAAATGCAAGAGAGTGGCTGTGTGATCTTACCGACTTCATAT 2238
QY 2333 TAAACTGCCCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAACCTTACATTTATCAGTC 2392
Db 2239 GCATGGCCCCAATCTCATTTCTATGCTCTGTAGGAATTTCTGAAAGCCTCTCATCACTG 2298
QY 2393 CTGAAGTAATTAAGTTTATCTCTTGTGGTGTAGTCCCACTTCTCTGCATGTCTCAATCCCC 2452
Db 2299 TTAGCAACTCCAAAATCTTGTGTGTACTTCTTCTATCCACTTAACTCTCTGTGCCAATCCAT 2358
```

```
QY 2453 TTCTCTACATCTGTTCAAATCTCTCACTTTAAGGAGGATCTGCTGAGCCT 2501
Db 2359 TCCTCTATGTATTTTTCACCAAGCCCTTCCAGAGGATGTGTTTCATCT 2407

RESULT 15
US-07-741-453A-57
; Sequence 57, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4417 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-07-741-453A-57
```

Query Match 3.7%; Score 100; DB 3; Length 4417;

Best Local Similarity 45.1%; Pred. No. 2.8e-19;
Matches 413; Conservative 0; Mismatches 500; Indels 3; Gaps 1;

```
QY 1616 TGCAGTGTTCACCTTCCCGAGGCCCTTCAAACCTGTGAACACCTGCTGTGATGGCTGGC 1675
Db 1249 TGGTGTGTACTCTTAAGTCAGATGAGTCAACCTCTGAAGACATATATGGCTACAAGT 1308
QY 1676 TGATCAGAAATGGAGTGTGGACCAATAGCAGTTCGTGGCACTTACTTGTAAATGCTTTGGTGA 1735
Db 1309 TCCTGAGGATTTGGTGTGGTGTGTTAGTCTGTCTGCTCTCTGGGCAATGCTTTGTGTC 1368
QY 1736 CTTCAACAGTTTTAGATCCCTCTGTATACATTTCCCAATTAAGTGTAAATGGGTGCA 1795
Db 1369 TGATCGTCTCTTACCAGTCACTACAATTTGACTGTGCCACGCTTTCTCATGTGCAACT 1428
QY 1796 TCGCAGCAGTGAACATGCTCAGCGAGTCTCCAGTGGCGTGTGGCTGGTGTGATGGT 1855
Db 1429 TGGCTTTGAGATTTCTGCATGGGATGTATCTGCTCTCATCGCTCCGCTAGACCTCT 1488
QY 1856 TCACTTTTGGCAGCTTTTGCAGCATGTTGCTGGTGGGAGAAATGGGTTGTGTCGANG 1915
Db 1489 ACACCTATCTGAGTACTTACAACCATGCAATCGATCGACTGGCAGACAGGCCCTGGGTGAACA 1548
```


This Page Blank (uspto)

[illegible]

Db 1 ATGGACACCTCCCGGCTCGGTGCTGCTGCTCTTGGCTGTGCTGCTGCGACGCTGGCGACC 60
Qy 61 GGGGGCAGCTCTCCAGAGTCTGGTGTGTGCTGAGGGGCTGCGCCACACACACTGTCATTGC 120
Db 61 GGGGGCAGCTCTCCAGAGTCTGGTGTGTGCTGAGGGGCTGCGCCACACACTGTCATTGC 120
Qy 121 GAGCCGACGGCAGGATGCTGCTCAGGGTGGACTGCTCCGACCTGGGCTCTCCGAGCTG 180
Db 121 GAGCCGACGGCAGGATGCTGCTCAGGGTGGACTGCTCCGACCTGGGCTCTCCGAGCTG 180
Qy 181 CTTTCCAACTCAGCGTCTTCACTCTACCTAGACCTCAGTATGAAACAAATCAGTCAG 240
Db 181 CTTTCCAACTCAGCGTCTTCACTCTACCTAGACCTCAGTATGAAACAAATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGAGGATACGTCCTTGGGGAAAC 300
Db 241 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTGAGAGGATACGTCCTTGGGGAAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Qy 361 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT 420
Qy 421 CAATCCCTGCGTCTGGATGCTAAACCAATCAGCTATGTGCGCCCAAGCTGTTTCAGTGGC 480
Db 421 CAATCCCTGCGTCTGGATGCTAAACCAATCAGCTATGTGCGCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTTGAGGACCTGCTGGCTGATGCAATGCTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTCCTTGAGGACCTGCTGGCTGATGCAATGCTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGTATGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGTATGGCTCCACAGCTTAGAGACTTTAGAT 720
Qy 721 TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAA 780
Db 721 TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTATAGCAACAATATCAGGTGCTACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTATAGCAACAATATCAGGTGCTACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATACTTCTATGACAAATCCATCCCAATTTGTTGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACTTCTATGACAAATCCATCCCAATTTGTTGGAGATCTGCT 900
Qy 901 TTTCAACAATTTAACCCTGAACCTAAGAACACTGACTCTGAAATGGTGCTCAAAATAACTGAA 960
Db 901 TTTCAACAATTTAACCCTGAACCTAAGAACACTGACTCTGAAATGGTGCTCAAAATAACTGAA 960
Qy 961 TTTCTCTGATTTAACCCTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
Db 961 TTTCTCTGATTTAACCCTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTTCAAGTCTAGATCTGCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTTCAAGTCTAGATCTGCT 1080
Qy 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC 1140
Db 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC 1140

Qy 1141 CTAAGACATTAATGAATCTAGGAATTAAGTTGACACTTTCCAGCAGTTGCTTTAGCCTC 1200
Db 1141 CTAAGACATTAATGAATCTAGGAATTAAGTTGACACTTTCCAGCAGTTGCTTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTACACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTACACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCTTTTCCCTTAACT 1320
Db 1261 TTGCCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCTTTTCCCTTAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAAACCTTTCCAGAACTCAAGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db 1381 TCATCTGAAAACCTTTCCAGAACTCAAGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Qy 1441 GCATTTGAGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAA 1500
Db 1441 GCATTTGAGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAA 1500
Qy 1501 AGCAGTATGGACGACCTTCAATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGT 1560
Db 1501 AGCAGTATGGACGACCTTCAATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGT 1560
Qy 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGAAAGACCTGAAAGCCCTTTCATTTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGAAAGACCTGAAAGCCCTTTCATTTCAGTGCAG 1620
Qy 1621 TGTTTCACTTTCCCGAGGCCCCCTTCAAAACCTGTGAAACACCTGCTGATGCTGATC 1680
Db 1621 TGTTTCACTTTCCCGAGGCCCCCTTCAAAACCTGTGAAACACCTGCTGATGCTGATC 1680
Qy 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTATGCTTTTGGTGCATTTCA 1740
Db 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTATGCTTTTGGTGCATTTCA 1740
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATTTGGGGTTCATCGCA 1800
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATTTGGGGTTCATCGCA 1800
Qy 1801 GCAGTGAACTATGCTCAGGGAGTCTCCAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 GCAGTGAACTATGCTCAGGGAGTCTCCAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 TTTGGCAGCTTTGACGACATGCTGCTGGGAGAAATGCGGTTGGTGGCAATGTCATT 1920
Db 1861 TTTGGCAGCTTTGACGACATGCTGCTGGGAGAAATGCGGTTGGTGGCAATGTCATT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCAAGAAATCACTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTTCAAGAAATCACTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCCAATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCCAATTTTCTAGC 2040
Qy 2041 CTGAAAGTAATCAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 CTGAAAGTAATCAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220
Db 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220

Db	241		CTGCTCCGGAATCCCTCTGCCAGTCTCCGCTTCTCTGGAGGAGTTAGCTCTTTCGGGAAAC	300
Qy	301	GCTCTGACATACATTTCCCAAGGAGCAATTCACATGGCCCTTTACAGTCTTAAAGTTCCTTATG	360	
Db	301	GCTCTGACATACATTTCCCAAGGAGCAATTCACATGGCCCTTTACAGTCTTAAAGTTCCTTATG	360	
Qy	361	CTGCAGAAATAATCAGCTTAAGACACGTACCCACAGAAGCTCTGCAGAAATTTGCGAAGCCTT	420	
Db	361	CTGCAGAAATAATCAGCTTAAGACACGTACCCACAGAAGCTCTGCAGAAATTTGCGAAGCCTT	420	
Qy	421	CAATCCCTGGTCTGGATGCTAACCAATCAGCTATGTGCCCCCGAAGCTGTTTCAGTGGC	480	
Db	421	CAATCCCTGGTCTGGATGCTAACCAATCAGCTATGTGCCCCCGAAGCTGTTTCAGTGGC	480	
Qy	481	CTGCATTTCCCTGAGGCACTGTGCTGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG	540	
Db	481	CTGCATTTCCCTGAGGCACTGTGCTGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG	540	
Qy	541	GCTTTTAGAAGTTTATCGGCAATTCGCAAGCCATGACCTTTGGCCCTGAAACAAATAACACCAC	600	
Db	541	GCTTTTAGAAGTTTATCGGCAATTCGCAAGCCATGACCTTTGGCCCTGAAACAAATAACACCAC	600	
Qy	601	ATACGACATATGCTTTTGGAAACCTCTCCAGCTTGTGTCTTCTACATCTCCATTAACAAT	660	
Db	601	ATACGACATATGCTTTTGGAAACCTCTCCAGCTTGTGTCTTCTACATCTCCATTAACAAT	660	
Qy	661	AGAAATCCATCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTTAGAT	720	
Db	661	AGAAATCCATCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTTAGAT	720	
Qy	721	TTAAATTAACAATAACCTTGTATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTTAAA	780	
Db	721	TTAAATTAACAATAACCTTGTATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTTAAA	780	
Qy	781	GAACTAGGATTTCTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840	
Db	781	GAACTAGGATTTCTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840	
Qy	841	CCCTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCAATTTGTGGGAGATCTGCT	900	
Db	841	CCCTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCAATTTGTGGGAGATCTGCT	900	
Qy	901	TTTCAACATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGCTCACAATAAATCTGAA	960	
Db	901	TTTCAACATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGCTCACAATAAATCTGAA	960	
Qy	961	TTTCTGATTTTAACTGGAACTGCAAACTCTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020	
Db	961	TTTCTGATTTTAACTGGAACTGCAAACTCTGGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020	
Qy	1021	TCACTCTCTCTCAACCGCTCGCAATCAGTTACTTAATCTCCAAGTGCTAGATCTGTCT	1080	
Db	1021	TCACTCTCTCTCAACCGCTCGCAATCAGTTACTTAATCTCCAAGTGCTAGATCTGTCT	1080	
Qy	1081	TACAACTTATTAGAAGATTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140	
Db	1081	TACAACTTATTAGAAGATTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140	
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAGTTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1200	
Db	1141	CTAAGACATAATGAAATCTACGAAATTTAAGTTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1200	
Qy	1201	CGATCGCTGAATTTGGCTTCGAAACAAATTTGCTATTATTACCCCCCATGCAATTTCCACT	1260	
Db	1201	CGATCGCTGAATTTGGCTTCGAAACAAATTTGCTATTATTACCCCCCATGCAATTTCCACT	1260	
Qy	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTGTCTTTTCTCTATAACT	1320	
Db	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTGTCTTTTCTCTATAACT	1320	
Qy	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAACAGGAAATCATGCCCTTACAGACTTGATA	1380	

QY 61 GGGGACGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCATTGC 120
Db 109 GGGGACGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCATTGC 168
QY 121 GAGCCGACGGCAGGATGTGCTCAGGGTGGACTGCTCCGACCTTGGGCTCTCGAGCTG 180
Db 169 GAGCCGACGGCAGGATGTGCTCAGGGTGGACTGCTCCGACCTTGGGCTCTCGAGCTG 228
QY 181 CTTTCCAACTCAGGCTCTTCCCTCCTACCTAGACCTCAGTATGAAACAATCAGTCAG 240
Db 229 CTTTCCAACTCAGGCTCTTCCCTCCTACCTAGACCTCAGTATGAAACAATCAGTCAG 288
QY 241 CTGCTCCGGAATCCCTCCGACGCTCCGCTTCTGAGAGGATGACGCTTGGGGAAC 300
Db 289 CTGCTCCGGAATCCCTCCGACGCTCCGCTTCTGAGAGGATGACGCTTGGGGAAC 348
QY 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTTCTTATG 360
Db 349 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTTAAAGTTCTTATG 408
QY 361 CTGCGAATTAATCAGCTAAGACAGTACCGACAGAACTCTGCGAATTTGGGAAGCCTT 420
Db 409 CTGCGAATTAATCAGCTAAGACAGTACCGACAGAACTCTGCGAATTTGGGAAGCCTT 468
QY 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 480
Db 469 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 528
QY 481 CTGCAATCCCTGAGGCACCTGTGGCTGGATGACCAATGCGTTTAAACAGAAATCCCGTCCAG 540
Db 529 CTGCAATCCCTGAGGCACCTGTGGCTGGATGACCAATGCGTTTAAACAGAAATCCCGTCCAG 588
QY 541 GCTTTTAGAAGTTTATCGGATGCAAGCATGACCTTGGCCCTGAAACAAATACACCAC 600
Db 589 GCTTTTAGAAGTTTATCGGATGCAAGCATGACCTTGGCCCTGAAACAAATACACCAC 648
QY 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGCTACATCTCCATAACAAT 660
Db 649 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGCTACATCTCCATAACAAT 708
QY 661 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCAAGCTTAGAGACTTTAGAT 720
Db 709 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCAAGCTTAGAGACTTTAGAT 768
QY 721 TTAATTAACAATAACCTTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db 769 TTAATTAACAATAACCTTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTAAA 828
QY 781 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 829 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 888
QY 841 CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Db 889 CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 948
QY 901 TTTGAACTTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
Db 949 TTTGAACTTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 1008
QY 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC 1020
Db 1009 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCACAGATC 1068
QY 1021 TCATCTCTTCTCAAACCTCTGCAATCAGTTTACCTTAATCTCCAAAGTCTAGATCTGTCT 1080
Db 1069 TCATCTCTTCTCAAACCTCTGCAATCAGTTTACCTTAATCTCCAAAGTCTAGATCTGTCT 1128
QY 1081 TACAACCTTATAGAGATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
Db 1129 TACAACCTTATAGAGATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1188
QY 1141 CTAAGACATTAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTTGTAGCCTC 1200

Db 1189 CTAAGACATTAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTTGCTTAGCCTC 1248
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTAATTAACCCCAATGCAATTTTCACT 1260
Db 1249 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTAATTAACCCCAATGCAATTTTCACT 1308
QY 1261 TTGCGCATCCCTAAATAAGCTGGACCTATCGTCCAACTCTGCTGCTCTTTTCCCTATAACT 1320
Db 1309 TTGCGCATCCCTAAATAAGCTGGACCTATCGTCCAACTCTGCTGCTCTTTTCCCTATAACT 1368
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1380
Db 1369 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1428
QY 1381 TCATCTGAAATCTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db 1429 TCATCTGAAATCTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGTGTCTGT 1488
QY 1441 GCATTTGAGTGTCTGAGAACTGCTTATAAGATTTCTCAATCAATGAAATTAAGGTGACAA 1500
Db 1489 GCATTTGAGTGTCTGAGAACTGCTTATAAGATTTCTCAATCAATGAAATTAAGGTGACAA 1548
QY 1501 AGCAGTATGGAACGCTTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTTCAAGATGAA 1560
Db 1549 AGCAGTATGGAACGCTTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTTCAAGATGAA 1608
QY 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1609 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1668
QY 1621 TGTTCACCTTCCAGAGCCCTTCAAACTGTCGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1669 TGTTCACCTTCCAGAGCCCTTCAAACTGTCGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 1728
QY 1681 AGAATTTGAGTGTGGACCATAGCACTTCTGSCACTTACTTGTAAATGCTTTTGGTGACTTCA 1740
Db 1729 AGAATTTGAGTGTGGACCATAGCACTTCTGSCACTTACTTGTAAATGCTTTTGGTGACTTCA 1788
QY 1741 ACAGTTTTCAGATCCCTCTGTGATATTTCCCCCAATTAACCTGTTAAATGGGGTCAATGCA 1800
Db 1789 ACAGTTTTCAGATCCCTCTGTGATATTTCCCCCAATTAACCTGTTAAATGGGGTCAATGCA 1848
QY 1801 GCAGTGAACATGCTCAGGGAGTCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1849 GCAGTGAACATGCTCAGGGAGTCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
QY 1861 TTTGGCAGCTTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1909 TTTGGCAGCTTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968
QY 1921 GGTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 1980
Db 1969 GGTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2028
QY 1981 GAGGTGGTGTCTGCTGAAATATTTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Db 2029 GAGGTGGTGTCTGCTGAAATATTTTGAAACGAAAGCTTCCATTTTCTAGC 2088
QY 2041 CTGAAGTAATTAATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2089 CTGAAGTAATTAATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
QY 2101 CTGGTGGCAGCAAGTATGGCCCTTCCCTCTCTGCTGCTGCTTGTGCTTGTGGGAGGCC 2160
Db 2149 CTGGTGGCAGCAAGTATGGCCCTTCCCTCTCTGCTGCTGCTTGTGCTTGTGGGAGGCC 2208
QY 2161 AGCACCATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 2209 AGCACCATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2268
QY 2221 ACCAATGCTTACACCAAGCTCTACTGCAATTTGGAACAAGGAGACCTTGGAGAAATTTTGG 2280

QY	421	CAATCCCTCGCTGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
DB	469	CAATCCCTCGCTGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	528
QY	481	CTGCATTCCTCAGGACCTGCTGGCTGGAATGCAATGCGTTAAACAGAAATCCCGTCCAG	540
DB	529	CTGCATTCCTCAGGACCTGCTGGCTGGAATGCAATGCGTTAAACAGAAATCCCGTCCAG	588
QY	541	GCTTTTGAAGTTTATCGGCAATTCGAAGCATGACCTTGGCCCTGAAACAAAATACACCAAC	600
DB	589	GCTTTTGAAGTTTATCGGCAATTCGAAGCATGACCTTGGCCCTGAAACAAAATACACCAAC	648
QY	601	ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATCAACAAT	660
DB	649	ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATCAACAAT	708
QY	661	AGAACTCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT	720
DB	709	AGAACTCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT	768
QY	721	TTAAATTTACAATAACCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAACTTAAA	780
DB	769	TTAAATTTACAATAACCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAACTTAAA	828
QY	781	GAACTAGGATTTATAGCAACAAATATCAGGTGATACCTTGAGAAAGCAATTTGTAGGCAAC	840
DB	829	GAACTAGGATTTATAGCAACAAATATCAGGTGATACCTTGAGAAAGCAATTTGTAGGCAAC	888
QY	841	CCCTCTCTTATTAACAATATTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT	900
DB	889	CCCTCTCTTATTAACAATATTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT	948
QY	901	TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA	960
DB	949	TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA	1008
QY	961	TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCAAGATC	1020
DB	1009	TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCAAGATC	1068
QY	1021	TCATCTCTCTCAACCGTCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGTCT	1080
DB	1069	TCATCTCTCTCAACCGTCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGTCT	1128
QY	1081	TACAACTTATTAAGAAATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC	1140
DB	1129	TACAACTTATTAAGAAATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC	1188
QY	1141	CTAAGACATATTAAGAAATTTACCAAGTTTAAAGTTGACATTTCCAGCAGTTGCTTAGCCCTC	1200
DB	1189	CTAAGACATATTAAGAAATTTACCAAGTTTAAAGTTGACATTTCCAGCAGTTGCTTAGCCCTC	1248
QY	1201	CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTTATTCACCCCAATGCAATTTCCACT	1260
DB	1249	CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTTATTCACCCCAATGCAATTTCCACT	1308
QY	1261	TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTAATCT	1320
DB	1309	TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTAATCT	1368
QY	1321	GGGTATCATGTTTAACTCACTTAAATTTAAACAGGAAATCATGCTTACAGAGCTTGATA	1380
DB	1369	GGGTATCATGTTTAACTCACTTAAATTTAAACAGGAAATCATGCTTACAGAGCTTGATA	1428
QY	1381	TCATCTGAAATTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGTGTCTGT	1440
DB	1429	TCATCTGAAATTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGTGTCTGT	1488
QY	1441	GCAATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1500
DB	1489	GCAATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1548

QY	1501	AGCAGTATGGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
DB	1549	AGCAGTATGGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1608
QY	1561	GACCTTTGAAGATTTCTGCTGCTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG	1620
DB	1609	GACCTTTGAAGATTTCTGCTGCTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG	1668
QY	1621	TGTTTCACTTCCCAAGGCCCTTCAAAACCTCTGTGAACACCTCTGTGAATGGCTGGCTGATC	1680
DB	1669	TGTTTCACTTCCCAAGGCCCTTCAAAACCTCTGTGAACACCTCTGTGAATGGCTGGCTGATC	1728
QY	1681	AGAAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA	1740
DB	1729	AGAAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA	1788
QY	1741	ACAGATTTTTCAGATCCCTCTGTACATTTTCCCCCACTTAAACTGTAAATGCGGTTCATCGCA	1800
DB	1789	ACAGATTTTTCAGATCCCTCTGTGTACATTTTCCCCCACTTAAACTGTAAATGCGGTTCATCGCA	1848
QY	1801	GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGTGCTGGCTGGTGTGGATGCGTTCATC	1860
DB	1849	GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGTGCTGGCTGGTGTGGATGCGTTCATC	1908
QY	1861	TTTGGCAGCTTTGCAAGCAATGCTGCTGGTGGGAGAAATGGGGTGGTGGATGCGTTCATC	1920
DB	1909	TTTGGCAGCTTTGCAAGCAATGCTGCTGGTGGGAGAAATGGGGTGGTGGATGCGTTCATC	1968
QY	1921	GGTGTGTGCTCCATTTTGTCTTCAAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
DB	1969	GGTGTGTGCTCCATTTTGTCTTCAAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2028
QY	1981	GAGCGTGGGTTCCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC	2040
DB	2029	GAGCGTGGGTTCCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC	2088
QY	2041	CTGAAAGTAAATCAATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2100
DB	2089	CTGAAAGTAAATCAATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2148
QY	2101	CTGGGTGGCAGCAAGTATGGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
DB	2149	CTGGGTGGCAGCAAGTATGGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2208
QY	2161	AGCACCATTGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCTCATGATG	2220
DB	2209	AGCACCATTGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCTCATGATG	2268
QY	2221	ACCATGCTTACACCAAGCTCTACTGCAATTTGGGAGAGAGCTGGAGAAATATTTGG	2280
DB	2269	ACCATGCTTACACCAAGCTCTACTGCAATTTGGGAGAGAGCTGGAGAAATATTTGG	2328
QY	2281	GACTGCTTATGTTAAACACATTTGCCCTGTTGCTTCTCACCACCTGCAATCTTAACTGC	2340
DB	2329	GACTGCTTATGTTAAACACATTTGCCCTGTTGCTTCTCACCACCTGCAATCTTAACTGC	2388
QY	2341	CCTGTGGCTTTCTGCTCTCTCTTCTTAAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2400
DB	2389	CCTGTGGCTTTCTGCTCTCTCTTCTTAAATAAACCCTTACATTTATCAGTCTCTGAAGTA	2448
QY	2401	ATTAAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2460
DB	2449	ATTAAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2508
QY	2461	ATCTGTTTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
DB	2509	ATCTGTTTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2568
QY	2521	TGGACAGATCAAAACACCAAGCTTGTGATGCTCAATTTAACTCTGTGATGCTCGAAAAACAG	2580
DB	2569	TGGACAGATCAAAACACCAAGCTTGTGATGCTCAATTTAACTCTGTGATGCTCGAAAAACAG	2628
QY	2581	TCCTGTGATCAACTCAAGCCTTGGTAAACCTTTTACCAGCTCCAGCATCATTATGACCTG	2640

Db 441 |||||CTGCTCCGGAATCCCTGCCAGTCTCGCTTCTTGAGGAGTACGTTTGGGGAAAC 500
Qy 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTTATG 360
Db 501 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTTATG 560
Qy 361 CTGCGAAGTAATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGGGAGGCCTT 420
Db 561 CTGCGAAGTAATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGGGAGGCCTT 620
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGCG 480
Db 621 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGCG 580
Qy 481 CTGCAATCCCTGAGGACACTGTGGCTGGATGACAAATGTGGTTAAAGAAATCCCGTCCAG 540
Db 681 CTGCAATCCCTGAGGACACTGTGGCTGGATGACAAATGTGGTTAAAGAAATCCCGTCCAG 740
Qy 541 GCTTTTAGAAGTTTATCGGCATTCAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
Db 741 GCTTTTAGAAGTTTATCGGCATTCAGCCATGACCTTGGCCCTGAAACAAATACACCAC 800
Qy 601 ATACCAGACTATGCTCTTTGGAAACCTCTCCAGCTTTGGTAGTTCTACATCTCCATAACAAT 660
Db 801 ATACCAGACTATGCTCTTTGGAAACCTCTCCAGCTTTGGTAGTTCTACATCTCCATAACAAT 860
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGTATGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 861 AGAATCCACTCCCTGGGAAAGAAATGCTTTGTATGATGGCTCCACAGCCTAGAGACTTTAGAT 920
Qy 721 TTAATATTACAAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAA 780
Db 921 TTAATATTACAAATAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAA 980
Qy 781 GAACTAGAGATTCATAGCAACAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 981 GAACTAGAGATTCATAGCAACAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 1040
Qy 841 CCTCTCTTATTAACAATATTCATGACAAATCCCAATCCCAATTTGTTGGGAGATCTGCT 900
Db 1041 CCTCTCTTATTAACAATATTCATGACAAATCCCAATCCCAATTTGTTGGGAGATCTGCT 1100
Qy 901 TTTCAACATTTACCTGAACATAGACACCTGACCTGCAATTTGGTGCCTCAACAATACGAA 960
Db 1101 TTTCAACATTTACCTGAACATAGACACCTGACCTGCAATTTGGTGCCTCAACAATACGAA 1160
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACCTTTAACTGGAGCACAGATC 1020
Db 1161 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACCTTTAACTGGAGCACAGATC 1220
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCAAGTGTAGATCTGTCT 1080
Db 1221 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCAAGTGTAGATCTGTCT 1280
Qy 1081 TACAACTTATAGAGATTTACCCAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1140
Db 1281 TACAACTTATAGAGATTTACCCAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1340
Qy 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACATCTTCCAGCAGTTGCTTAGGCCTC 1200
Db 1341 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACATCTTCCAGCAGTTGCTTAGGCCTC 1400
Qy 1201 CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTAATTAATCCACCCAAATGCAATTTCCACT 1260
Db 1401 CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTAATTAATTCACCCCAATGCAATTTCCACT 1460
Qy 1261 TTGGCATCCCTAAATAAGCTGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1320
Db 1461 TTGGCATCCCTAAATAAGCTGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACT 1520
Qy 1321 GGGTTACATGGTTAACTCACTTAAATTTAACAGGAAATCATGCTTACAGAGCTTGATA 1380

Db 1521 GGGTTACATGGTTAACTCACTTAAATTTAACAGGAAATCATGCTTACAGAGCTTGATA 1580
Qy 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACAGTGCTGT 1440
Db 1581 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACAGTGCTGT 1640
Qy 1441 GCATTTGAGTGTGTGAGAAATGCCTTATAAGATTTCTAATCAATGAAATAAAGGTGCAAC 1500
Db 1641 GCATTTGAGTGTGTGAGAAATGCCTTATAAGATTTCTAATCAATGAAATAAAGGTGCAAC 1700
Qy 1501 AGCAGTATGAGACGACCTTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAA 1560
Db 1701 AGCAGTATGAGACGACCTTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAA 1760
Qy 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGAAAGACCTGAAAGCCCTTCAATTCAGTGAG 1620
Db 1761 GACCTTGAAGATTTTCTGCTTGAATTTGAGAAAGACCTGAAAGCCCTTCAATTCAGTGAG 1820
Qy 1621 TGTTTCACTTCCCAAGGCCCTTCAAAACCTGTGAAACCTGCTTGATGGCTGGCTGATC 1680
Db 1821 TGTTTCACTTCCCAAGGCCCTTCAAAACCTGTGAAACCTGCTTGATGGCTGGCTGATC 1880
Qy 1681 AGAATTTGAGTGTGGAACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
Db 1881 AGAATTTGAGTGTGGAACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1940
Qy 1741 ACAGTTTTTCAATCCCTCTGTACATTTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATCGCA 1800
Db 1941 ACAGTTTTTCAATCCCTCTGTACATTTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATCGCA 2000
Qy 1801 CGAGTGAACATGCTCAACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGCGTTCAT 1860
Db 2001 CGAGTGAACATGCTCAACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGCGTTCAT 2060
Qy 1861 TTTGSCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGGTTCATGCTCAAT 1920
Db 2061 TTTGSCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGGTTCATGCTCAAT 2120
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCACTGTTTCTGCTTACCTCGCAGCCCTG 1980
Db 2121 GGTTTTTTGTCCATTTTGTCTTCAGAAATCACTGTTTCTGCTTACCTCGCAGCCCTG 2180
Qy 1981 GAGCGTGGTGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 2181 GAGCGTGGTGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2240
Qy 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCTTGAACATGCGCGCAGTTCCCTG 2100
Db 2241 CTGAAAGTAAATCAATTTTGTCTCTGTGCTTGAACATGCGCGCAGTTCCCTG 2300
Qy 2101 CTGGTGGCAGCAGATAGCGGCTCCCTCTCTGCTGCTGCTTGGGAGGCC 2160
Db 2301 CTGGTGGCAGCAGATAGCGGCTCCCTCTCTGCTGCTGCTTGGGAGGCC 2360
Qy 2161 AGCACATGGGCTACATGGTCTCTCATCTTGTCTCAAATTCCTTGTCTTCTCATGATG 2220
Db 2361 AGCACATGGGCTACATGGTCTCTCATCTTGTCTCAAATTCCTTGTCTTCTCATGATG 2420
Qy 2221 ACCATGCTTACACCAAGCTCTATGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2280
Db 2421 ACCATGCTTACACCAAGCTCTATGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2480
Qy 2281 GACTGCTTATGGTAAACACATGGCCTGCTGCTTTCACCAACCTGATCTCTAACTGC 2340
Db 2481 GACTGCTTATGGTAAACACATGGCCTGCTGCTTTCACCAACCTGATCTCTAACTGC 2540
Qy 2341 CCTGTGGCTTCTTGTCTCTCTCTTCTTAAATAAACTTATACATTTATCAGTCTCGAAGTA 2400
Db 2541 CCTGTGGCTTCTTGTCTCTCTCTTCTTAAATAAACTTATACATTTATCAGTCTCGAAGTA 2600
Qy 2401 ATTAAGTTTATCTTCTGCTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Db 2601 ATTAAGTTTATCTTCTGCTGGTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2660

```
QY 2461 ATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGGAGCTCGAAGAAAGCAACCTACGTC 2520
Db |||||
QY 2661 ATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGGAGCTCGAAGAAAGCAACCTACGTC 2720
Db |||||
QY 2521 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAAACAG 2580
Db |||||
QY 2721 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAAACAG 2780
Db |||||
QY 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTTATGACCTG 2640
Db |||||
QY 2781 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTTATGACCTG 2840
Db |||||
QY 2641 CTCTCCAGTTCGTCGCATCACAGCTTATCCAGTGA CTGAGAGCTGCCATCTTCTCTCT 2700
Db |||||
QY 2841 CTCTCCAGTTCGTCGCATCACAGCTTATCCAGTGA CTGAGAGCTGCCATCTTCTCTCT 2900
Db |||||
QY 2701 GTGGCATTTGTCCCATGTCTCTAA 2724
Db |||||
QY 2901 GTGGCATTTGTCCCATGTCTCTAA 2924
Db |||||
```

RESULT 10

```
US-10-295-027-945
; Sequence 945, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 945
; LENGTH: 3032
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-945
```

Query Match 100.0%; Score 2724; DB 17; Length 3032;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACACCTCCCGCTCGGTGTCTCTGTCTCTTGCCTGTGCTGCTGAGTCGACGACC 60
Db |||||
QY 201 ATGGACACCTCCCGCTCGGTGTCTCTGTCTCTTGCCTGTGCTGAGTCGACGACC 260
Db |||||
QY 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCACACACCTGTCTATGCG 120
Db |||||
QY 261 GGGGGCAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCACACACCTGTCTATGCG 320
Db |||||
QY 121 GAGCCCGACGGCAGGATGTGCTCAGGGGTGGAAGCTCTCCGACTCTGGGGCTCTCGGAGCTG 180
Db |||||
QY 321 GAGCCCGACGGCAGGATGTGCTCAGGGGTGGAAGCTCTCCGACTCTGGGGCTCTCGGAGCTG 380
Db |||||
QY 181 CCTTCCAACTCAGGCTCTTCACTCTCTTACCTAGACCTCAGTATGAA CAACATCAGTCAAG 240
Db |||||
QY 381 CCTTCCAACTCAGGCTCTTCACTCTCTTACCTAGACCTCAGTATGAA CAACATCAGTCAAG 440
Db |||||
QY 241 CTGCTCCCGAATCCCTGCGCTCTCCGCTCTCTGGAGGAGTTACGCTTTGGGGGAAC 300
Db |||||
QY 441 CTGCTCCCGAATCCCTGCGCTCTCCGCTCTCTGGAGGAGTTACGCTTTGGGGGAAC 500
Db |||||
QY 301 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTTACGCTCTTAAAGTCTTATG 360
Db |||||
QY 501 GCTCTGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTTACGCTCTTAAAGTCTTATG 560
Db |||||
QY 361 CTGCAGAAATAATCAGCTAAGACACGTACCCACAGAAAGCTCTGCAGAAATTTGGGAAGCTTT 420
Db |||||
QY 561 CTGCAGAAATAATCAGCTAAGACACGTACCCACAGAAAGCTCTGCAGAAATTTGGGAAGCTTT 620
Db |||||
QY 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTTTTCAAGTGGC 480
Db |||||
QY 621 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTTTTCAAGTGGC 680
Db |||||
QY 481 CTGCATTTCCCTGAGGCACCTGTGGCTGATGACATGGTTAAACAGAAATCCCGCTCCAG 540
Db |||||
QY 681 CTGCATTTCCCTGAGGCACCTGTGGCTGATGACATGGTTAAACAGAAATCCCGCTCCAG 740
Db |||||
QY 541 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAAATACACCA 600
Db |||||
QY 741 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAAATACACCA 800
Db |||||
QY 601 ATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTGTGTTCTACATCTCCATAACAT 860
Db |||||
QY 801 ATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTGTGTTCTACATCTCCATAACAT 860
Db |||||
QY 661 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTCAGAGACTTTAGAT 920
Db |||||
QY 861 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTCAGAGACTTTAGAT 920
Db |||||
QY 721 TTAATAATTACAATAAACCTTTGATGAAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAA 780
Db |||||
QY 921 TTAATAATTACAATAAACCTTTGATGAAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAA 980
Db |||||
QY 781 GAATAGGATTTATAGCAACAAATATCAGGTGCTGATACCTGAGAAAGCAATTTGAGGCAAC 840
Db |||||
QY 981 GAATAGGATTTATAGCAACAAATATCAGGTGCTGATACCTGAGAAAGCAATTTGAGGCAAC 1040
Db |||||
QY 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCATTTGTTGGGAGATCTGCT 900
Db |||||
QY 1041 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCATTTGTTGGGAGATCTGCT 1100
Db |||||
QY 901 TTTTCAACATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGCCTCACAATAACTGAA 960
Db |||||
QY 1101 TTTTCAACATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGCCTCACAATAACTGAA 1160
Db |||||
QY 961 TTTTCTGATTTTAACTGGAACTGGAAACCTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
Db |||||
QY 1161 TTTTCTGATTTTAACTGGAACTGGAAACCTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1220
Db |||||
QY 1021 TCATCTCTTCTCAACACCTGCTGCAATCAGTTTACCTTAATCTCCAAGTGTAGATCTGCT 1080
Db |||||
QY 1221 TCATCTCTTCTCAACACCTGCTGCAATCAGTTTACCTTAATCTCCAAGTGTAGATCTGCT 1280
Db |||||
```


QY 1261 TTGGCATCCCTAATAAGAGCTGGACCTATCGTCCAACTCCTCGTGGTCTTTCTTAACT 1320
Db 1261 TTGGCATCCCTAATAAGAGCTGGACCTATCGTCCAACTCCTCGTGGTCTTTCTTAACT 1320
QY 1321 GGGTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCCCTTACAGAGCTTGATA 1380
Db 1321 GGGTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCCCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAAACTTTTCGAGAACTCAAGGTTATAGAAATGCTTAACTCAAGTGTGT 1440
Db 1381 TCATCTGAAAACTTTTCGAGAACTCAAGGTTATAGAAATGCTTAACTCAAGTGTGT 1440
QY 1441 GCATTTGGAGTGTGAGAAATGCTTAACTCAAGTGTGTGAGAAATGCTTAACTCAAGTGTGT 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTAACTCAAGTGTGTGAGAAATGCTTAACTCAAGTGTGT 1500
QY 1501 AGCAGTATGGAAGACCTTCAAGAAAGATGCTGGAATGCTTCAAGTGTGTGAGAAATGCTTAACTCAAGTGTGT 1560
Db 1501 AGCAGTATGGAAGACCTTCAAGAAAGATGCTGGAATGCTTCAAGTGTGTGAGAAATGCTTAACTCAAGTGTGT 1560
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGTGT 1620
Db 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGTGT 1620
QY 1621 TGCTTCACTTCCCGAGCCCTTCAAGCCCTGTAACCTGCTGAGTGTGTGAGTGTGTGAGTGTGT 1680
Db 1621 TGCTTCACTTCCCGAGCCCTTCAAGCCCTGTAACCTGCTGAGTGTGTGAGTGTGTGAGTGTGT 1680
QY 1681 AGAATGGAGTGTGAGACCTAGCAGTCTGGACCTTACCTTGAATGCTTGGTGTGTGTGTGTGT 1740
Db 1681 AGAATGGAGTGTGAGACCTAGCAGTCTGGACCTTACCTTGAATGCTTGGTGTGTGTGTGTGTGT 1740
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAATGGGGTGTGTGTGTGTGTGT 1800
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAATGGGGTGTGTGTGTGTGTGTGT 1800
QY 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGGCTGCTGGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGGCTGCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
QY 1861 TTTGGCAGCTTTGACAGCATGCTGCTGGTGGGAGATGGGTTGGTGGTGGTGGTGGTGGTGGTGGT 1920
Db 1861 TTTGGCAGCTTTGACAGCATGCTGCTGGTGGGAGATGGGTTGGTGGTGGTGGTGGTGGTGGTGGT 1920
QY 1921 GGTGTTTTCAGATTTTGTCTTCAAGATCATCTGTTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCT 1980
Db 1921 GGTGTTTTCAGATTTTGTCTTCAAGATCATCTGTTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCT 1980
QY 1981 GAGCGTGGTCTCTGTGAAATATTTCTGAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTCTGAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAACTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 CTGAAAGTAACTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 CTGGGTGGCAGCAAGTATGGGCTTCCCTCTCTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2160
Db 2101 CTGGGTGGCAGCAAGTATGGGCTTCCCTCTCTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2160
QY 2161 AGCAGTGGGCTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 AGCAGTGGGCTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGAAGAGGAGCTGGAGATTTGG 2280
Db 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGAAGAGGAGCTGGAGATTTGG 2280
QY 2281 GACTGCTTATGGAAGAAACATTTGCTGCTTCTTCAACCACTGCACTTCAACTGC 2340
Db 2281 GACTGCTTATGGAAGAAACATTTGCTGCTTCTTCAACCACTGCACTTCAACTGC 2340
QY 2341 CCTGTGGCTTTCTGTCTTCTCTCTTTTAAATAAACCTTATCATGCTGCAAGTA 2400

Db 2341 CCTGTGGCTTTCTGTGCTTCTCTCTTTTAAATAAACCTTATCATGCTGCAAGTA 2400
QY 2401 AATAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 2401 AATAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
QY 2461 ATCTGTTTCAATCTTCACTTTAAGGAGGATCTGGTGGGCTGAGAAAGCAAACTTACGTC 2520
Db 2461 ATCTGTTTCAATCTTCACTTTAAGGAGGATCTGGTGGGCTGAGAAAGCAAACTTACGTC 2520
QY 2521 TGGAACAAGATCAAAACACCCAGCTTGTGCTCAATTAACCTGATGATGCTGATGATGCTGATGATGCTGATGATGCT 2580
Db 2521 TGGAACAAGATCAAAACACCCAGCTTGTGCTCAATTAACCTGATGATGCTGATGATGCTGATGATGCTGATGATGCT 2580
QY 2581 TCCTGTGATCACTCAAGCTTGTGCTTACAGCTTCCAGCTTCACTTATGACCTG 2640
Db 2581 TCCTGTGATCACTCAAGCTTGTGCTTACAGCTTCCAGCTTCACTTATGACCTG 2640
QY 2641 CTCCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 CTCCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
QY 2701 GTGGCAATTTGCTCCATGCTC 2721
Db 2701 GTGGCAATTTGCTCCATGCTC 2721

RESULT 13

US-10-295-027-483

; Sequence 483, Application US/10295027

; Publication No. US20030232350A1

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

; FILE REFERENCE: 018501-012500US

; CURRENT APPLICATION NUMBER: US/10/295,027

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: US 09/663,733

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/335,394

; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464

; PRIOR FILING DATE: 2001-11-21

; PRIOR APPLICATION NUMBER: US 60/334,393

; PRIOR FILING DATE: 2001-11-29

; PRIOR APPLICATION NUMBER: US 60/340,376

; PRIOR FILING DATE: 2001-12-14

; PRIOR APPLICATION NUMBER: US 60/347,211

; PRIOR FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US 60/347,349

; PRIOR FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/355,250

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: US 60/356,714

; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1386

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 483

; LENGTH: 2651

Db 661 AGAATCCATCCCTCGGAGAAAGAAATGCTTTTGATGGCTCCACAGCTAGAGACTTTAGAT 720
Qy 721 TTAATTTACAATAAACCCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAACTTTAAA 780
Db 721 TTAATTTACAATAAACCCTTGATGAATTTCCCACTGCAATTTAGACACTCTCCAACTTTAAA 780
Qy 781 GAACATAGGATTTTCATAGCAACAATATCAGTGCATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 781 GNA----- 784
Qy 841 CTTCTCTTTATTAACAATATTTATGACAAATCCCAATTTGTTGGAGATCTGCT 900
Db 785 -----TACATTTCTATGACAATCCCATCCAATTTGTTGGAGATCTGCT 828
Qy 901 TTTCAACATTTACCTGAACCTAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
Db 829 TTTCAACATTTACCTGAACCTAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 888
Qy 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 889 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 948
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGCT 1080
Db 949 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAAGTGTAGATCTGCT 1008
Qy 1081 TACAACCTATTAGAAGATTTACCAAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
Db 1009 TACAACCTATTAGAAGATTTACCAAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1068
Qy 1141 CTAAGACATAATGAATCTACGAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Db 1069 CTAAGACATAATGAATCTACGAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1128
Qy 1201 CGATCGCTGAATTTGGCTGGAAACAAAATGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 1129 CGATCGCTGAATTTGGCTGGAAACAAAATGCTATTATTCACCCCAATGCAATTTCCACT 1188
Qy 1261 TTGCCATCGCTAATAAGCTGACACTATGCTTCAACCTCTGCTCTTTTCCCTATAACT 1320
Db 1189 TTGCCATCGCTAATAAGCTGACACTATGCTTCAACCTCTGCTCTTTTCCCTATAACT 1248
Qy 1321 GGGTTACATGGTTAACTCAGTTAACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 1249 GGGTTACATGGTTAACTCAGTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1308
Qy 1381 TCATCTGAATACTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACAGTGTGT 1440
Db 1309 TCATCTGAATACTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACAGTGTGT 1368
Qy 1441 GCATTTGGAGTGTGAGNAATGCCTATAAGATTTCTAATCATGGNAATAAGGTGACAAAC 1500
Db 1369 GCATTTGGAGTGTGAGNAATGCCTATAAGATTTCTAATCATGGNAATAAGGTGACAAAC 1428
Qy 1501 AGCAGTATGAGAGCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACT 1560
Db 1429 AGCAGTATGAGAGCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACT 1488
Qy 1561 GACCTTGAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1489 GACCTTGAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1548
Qy 1621 TGTTCACTTCCAGGCCCTTCAACCCCTGTAACACCTGCTGTGATGGCTGGCTGATC 1680
Db 1549 TGTTCACTTCCAGGCCCTTCAACCCCTGTAACACCTGCTGTGATGGCTGGCTGATC 1608
Qy 1681 AGAATTTGGAGTGTGGACCAATAGACTTCTGGCACTTACTTTGTAATGCTTTGTGACTTCA 1740
Db 1609 AGAATTTGGAGTGTGGACCAATAGACTTCTGGCACTTACTTTGTAATGCTTTGTGACTTCA 1668
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1800
Db 1669 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA 1728

Qy 1801 GCAGTGAACATGCTTCCACGGAGTCTCCAGTGCCTGGCTGGTGGATGCTGATCGTTCACT 1860
Db 1729 GCAGTGAACATGCTTCCACGGAGTCTCCAGTGCCTGGCTGGTGGATGCTGATCGTTCACT 1788
Qy 1861 TTTGGCAGCTTTTGACAGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTCATTCATT 1920
Db 1789 TTTGGCAGCTTTTGACAGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTCATTCATT 1848
Qy 1921 GGTTTTTTGTGCAATTTTGTGCTCAGAAATCATCTGTTTCTGCTTACTCTGCGACGCCCTG 1980
Db 1849 GGTTTTTTGTGCAATTTTGTGCTCAGAAATCATCTGTTTCTGCTTACTCTGCGACGCCCTG 1908
Qy 1981 GAGCGTGGTGTCTCTGTGCAAAATATTCTGCAAAATTTTGAAGCAAGCTCCATTTTCTAGC 2040
Db 1909 GAGCGTGGTGTCTCTGTGCAAAATATTCTGCAAAATTTTGAAGCAAGCTCCATTTTCTAGC 1968
Qy 2041 CTGAAGTAATCATTTTGTCTCTGTGCGCTGTGCGCTTTGACCATGGCCGAGTTCGCCCTG 2100
Db 1969 CTGAAGTAATCATTTTGTCTCTGTGCGCTGTGCGCTTTGACCATGGCCGAGTTCGCCCTG 2028
Qy 2101 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTGCTTTGCGCTTTTGGGGAGCCCC 2160
Db 2029 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTGCTTTGCGCTTTTGGGGAGCCCC 2088
Qy 2161 AGCACCATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 2089 AGCACCATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2148
Qy 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACTGGAGAATATTTGG 2280
Db 2149 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACTGGAGAATATTTGG 2208
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCAACCACTGATCCTTAACCTGC 2340
Db 2209 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCAACCACTGATCCTTAACCTGC 2268
Qy 2341 CCTGGGCTTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTCAAGTA 2400
Db 2269 CCTGGGCTTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTCAAGTA 2328
Qy 2401 ATTAAGTTTATCTCTTCTGGTGGTAGTCCCACCTTCTGCAATGCTCAATCCCCTTCTCTAC 2460
Db 2329 ATTAAGTTTATCTCTTCTGGTGGTAGTCCCACCTTCTGCAATGCTCAATCCCCTTCTCTAC 2388
Qy 2461 ATCTTGTTCATCTCTCATTAAAGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2520
Db 2389 ATCTTGTTCATCTCTCATTAAAGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2448
Qy 2521 TGGACAAGATCAAAACACCCCAAGCTTGTATGTCATTAATACTCTGATGATGTCGAAAAACAG 2580
Db 2449 TGGACAAGATCAAAACACCCCAAGCTTGTATGTCATTAATACTCTGATGATGTCGAAAAACAG 2508
Qy 2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db 2509 TCCTGTGACTCAACTCAAGCCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2568
Qy 2641 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTGTGAGAGCTGCCATCTTCTCTCT 2700
Db 2569 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTGTGAGAGCTGCCATCTTCTCTCT 2628
Qy 2701 GTGGCAATTTGTCCCCATGCT 2720
Db 2629 GTGGCAATTTGTCCCCATGCT 2648

RESULT 15

US-10-173-999-27
; Sequence 27, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.

Db	661	AGAA	TCACTCCCTCGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Qy	721	TTAAAT	TACAATAAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCTCCAACTTTAAA	780
Db	721	TTAAAT	TACAATAAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCTCCAACTTTAAA	780
Qy	781	GAAC	TAGGATTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTTAGGCAC	840
Db	781	GAAC	-----TAAAC-----	784
Qy	841	CCTTCTCT	TATACAAATACATTTTCTATGACAAATCCCATCCAAATTTGTTGGGAGACTTGCT	900
Db	785	-----	-----TACATTTCTATGACATCCCATCCAAATTTGTTGGGAGACTTGCT	828
Qy	901	TTTTCAAC	ATTTACTGAACTAAGAACACACTGACTCTGAAATGSGTCCCTCACAATAA	960
Db	829	TTTTCAAC	ATTTACTGAACTAAGAACACACTGACTCTGAAATGSGTCCCTCACAATAA	888
Qy	961	TTTTCTCTG	ATTTAACTGGAACCTGAAACCTGGAGAGCTGACTTTTA	1020
Db	889	TTTTCTCTG	ATTTAACTGGAACCTGAAACCTGGAGAGCTGACTTTTA	948
Qy	1021	TCACTCT	TTCTCCAAACCGTCTGCAATCAGTTACCTAACTCTCCAAGTGTCTAGATCTGCT	1080
Db	949	TCACTCT	TTCTCCAAACCGTCTGCAATCAGTTACCTAACTCTCCAAGTGTCTAGATCTGCT	1008
Qy	1081	TACAAC	CTATTTAGAAGATTTACCACAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAAA	1140
Db	1009	TACAAC	CTATTTAGAAGATTTACCACAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAAA	1068
Qy	1141	CTAAGA	CATAATGAAATCTACGAAATTTAAAGTTTGAACATTTTCAGCAGTTTGC	1200
Db	1069	CTAAGA	CATAATGAAATCTACGAAATTTAAAGTTTGAACATTTTCAGCAGTTTGC	1128
Qy	1201	CGATCCGT	GTGAATTTGGCTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT	1260
Db	1129	CGATCCGT	GTGAATTTGGCTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT	1188
Qy	1261	TTGCCAT	CCCTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCGCTTTTCCTCTAACT	1320
Db	1189	TTGCCAT	CCCTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCGCTTTTCCTCTAACT	1248
Qy	1321	GGGTTC	CATGGTTTAACTCACTTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1249	GGGTTC	CATGGTTTAACTCACTTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1308
Qy	1381	TCATCTG	AAAACTTTCCAGAACCTCAAGGTTATAGAAATGGCTTATGCTTTACCAAGTCTGT	1440
Db	1309	TCATCTG	AAAACTTTCCAGAACCTCAAGGTTATAGAAATGGCTTATGCTTTACCAAGTCTGT	1368
Qy	1441	GCATTTGG	AGTGTGTCAGAGATGCTATAAGATTTCTAATCAATGGAATTAAGGTGACAACT	1500
Db	1369	GCATTTGG	AGTGTGTCAGAGATGCTATAAGATTTCTAATCAATGGAATTAAGGTGACAACT	1428
Qy	1501	AGCAGT	ATGACGACCTTCATAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1429	AGCAGT	ATGACGACCTTCATAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1488
Qy	1561	GACCTTG	AGATTTCTGCTTTGACTTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG	1620
Db	1489	GACCTTG	AGATTTCTGCTTTGACTTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG	1548
Qy	1621	TGTTCACT	TTCCCGAGCCCTTTCAAACTCTGGAACACTGCTTCATGGCTGGCTGATC	1680
Db	1549	TGTTCACT	TTCCCGAGCCCTTTCAAACTCTGGAACACTGCTTCATGGCTGGCTGATC	1608
Qy	1681	AGAA	TTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1609	AGAA	TTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1668
Qy	1741	ACAG	TTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACATGTTAAATGGGGTCATCGA	1800
Db	1669	ACAG	TTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACATGTTAAATGGGGTCATCGA	1728

Qy	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACT	1860
Db	1729	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACT	1788
Qy	1861	TTTGGCAGCTTTGACACGACATGGTGCCTGCTGGTGGGAGATGGGGTTGGTGCATGTCAAT	1920
Db	1789	TTTGGCAGCTTTGACACGACATGGTGCCTGCTGGTGGGAGATGGGGTTGGTGCATGTCAAT	1848
Qy	1921	GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACCTCTGGCAGCCCTG	1980
Db	1849	GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACCTCTGGCAGCCCTG	1908
Qy	1981	GAGCGTGGTCTCTGTGAAATATTTCTGAAATATTTCTGAAATATTTGAAACGAAAGCTCCCAATTTCTAGC	2040
Db	1909	GAGCGTGGTCTCTGTGAAATATTTCTGAAATATTTCTGAAATATTTGAAACGAAAGCTCCCAATTTCTAGC	1968
Qy	2041	CTGAAAGTAATCATTTTCTCTCTGTGCCCTGTGCTGCTGACCATGGCCGACAGTTCCCTCTG	2100
Db	1969	CTGAAAGTAATCATTTTCTCTCTGTGCCCTGTGCTGCTGACCATGGCCGACAGTTCCCTCTG	2028
Qy	2101	CTGGGTGACAGAGATATGGGCTCTCCCTCTCTGCTGCTGCTTTGGGGAGCCC	2160
Db	2029	CTGGGTGACAGAGATATGGGCTCTCCCTCTCTGCTGCTGCTTTGGGGAGCCC	2088
Qy	2161	AGCACCATGGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG	2220
Db	2089	AGCACCATGGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG	2148
Qy	2221	ACCATTTGCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATATTTGG	2280
Db	2149	ACCATTTGCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGACCTGGAGAAATATTTGG	2208
Qy	2281	GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCAACAACTGCATCCTAAACTGC	2340
Db	2209	GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCAACAACTGCATCCTAAACTGC	2268
Qy	2341	CCTGTGGCTTTCTGTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA	2400
Db	2269	CCTGTGGCTTTCTGTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA	2328
Qy	2401	ATTAAGTTTATCTTTCTGGTGGTAGTCCCACTTCTGCAATGTCTCAATCCCTTCTCTAC	2460
Db	2329	ATTAAGTTTATCTTTCTGGTGGTAGTCCCACTTCTGCAATGTCTCAATCCCTTCTCTAC	2388
Qy	2461	ATCTTGTTCATCTCTCACTTTAAGAGGATCTGGTGGCTGAGAAAGCAACCTACGTC	2520
Db	2389	ATCTTGTTCATCTCTCACTTTAAGAGGATCTGGTGGCTGAGAAAGCAACCTACGTC	2448
Qy	2521	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAATCTGTATGATGTGAAAAACAG	2580
Db	2449	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAATCTGTATGATGTGAAAAACAG	2508
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTG	2640
Db	2509	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTG	2568
Qy	2641	CCTCCAGTTCGGTGCATCAACAGCTTATCGAGTACTGAGAGCTGCCATCTTTCCTCT	2700
Db	2569	CCTCCAGTTCGGTGCATCAACAGCTTATCGAGTACTGAGAGCTGCCATCTTTCCTCT	2628
Qy	2701	GTGGCATTTGTCCCATGTCT	2720
Db	2629	GTGGCATTTGTCCCATGTCT	2648

Search completed: July 12, 2005, 21:05:45
Job time : 1831.02 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2652	97.4	72.24	9	AY411732	Homo sapi	
2	2004.6	73.6	27.24	9	AY411734	Mus muscu	
3	1958	71.9	25.93	9	AY411733	Pan trogl	
4	1037	38.1	11.04	1	AL530798	AL530798	
5	979.2	35.9	10.60	1	AL541959	AL541959	
6	845.8	31.6	47.96	3	AK047873	Mus muscu	
7	805.8	29.0	35.83	3	AK085901	Mus muscu	
8	745.2	27.1	9.40	7	CK449409	CK449409	
9	738.6	27.1	8.00	4	BG196708	BG196708	
10	669.4	24.6	6.71	4	BM795259	K-BST0077	
11	661.4	24.3	35.09	3	BC038795	Homo sapi	
12	604.4	22.2	7.63	7	CF742674	UI-M-HB0-	
13	585.4	21.5	8.73	5	BX403219	BX403219	
14	576.4	21.2	30.35	3	AK044357	Mus muscu	
15	573.4	21.0	5.90	6	CD517354	CD517354	
16	524	19.2	28.22	3	AK052873	Mus muscu	
17	521.2	19.1	7.76	5	BU0708439	UI-M-FC0-	
18	500.8	18.4	53.9	1	AI697103	AI697103	
19	486.4	17.9	6.95	6	BY731701	BY731701	
20	484.2	17.8	6.34	7	CF177761	CF177761	
21	475	17.4	4.75	5	BX107244	BX107244	
22	472	17.3	4.72	1	AA460529	AA460529	
23	471	17.3	4.94	1	AA424098	AA424098	
24	467.4	17.2	5.56	5	BQ323949	IL5-CI014	

Db 1 ATGGACACCTCCCGGCTCGGTGTGCTCCTGTCTTGTCCCTGTGCTGCTGCAGCTGGCGACC 60
Qy 61 GGGGGCAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATGCG 120
Db 61 GGGGGCAGCTCTCCAGGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATGCG 120
Qy 121 GAGCCCGAGCGGAGGATGTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Db 121 GAGCCCGAGCGGAGGATGTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Qy 181 CTTTCCAACTCAGCGTCTTCACTCTCTACCTCTAGACCTCAGTATGAACAAATCAGTCTG 240
Db 181 CTTTCCAACTCAGCGTCTTCACTCTCTACCTCTAGACCTCAGTATGAACAAATCAGTCTG 240
Qy 241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTTGCGGGAAAC 300
Db 241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTTGCGGGAAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGGAGCATTCAGTGGCCCTTTACAGTCTTTAAAGTCTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTCAGTGGCCCTTTACAGTCTTTAAAGTCTTNNN 360
Qy 361 CTGCAGAAATAACAGCTAAGACACGTAACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 NNN 420
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 NNNNNNNNGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTTGAGGACACTGTGGCTGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTCCTTGAGGACACTGTGGCTGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCATTCGACCAATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTGAAGTTTATCGGCATTCGACCAATGACCTTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGTCTACATCTCCATACAAAT 660
Db 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTGTGTCTACATCTCCATACAAAT 660
Qy 661 AGAATCCACTCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Qy 721 TTAATAATCAATTAACCTTTGATGAATTCGCCCACTGCAATTAGAGACACTCTCCAAACCTTAA 780
Db 721 TTAATAATCAATTAACCTTTGATGAATTCGCCCACTGCAATTAGAGACACTCTCCAAACCTTAA 780
Qy 781 GAACTAGGATTTTATAGCAACAAATATCAGTTCGATACCTGAGAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTTATAGCAACAAATATCAGTTCGATACCTGAGAGCAATTTGTAGGCAAC 840
Qy 841 CTTCTCTCTTATTAACAATATCTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CTTCTCTCTTATTAACAATATCTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACATTTA CTTGAA CTAAGAA CACTGACTCTGAATGGTGCCTCAAAATACTGAA 960
Db 901 TTTCAACATTTA CTTGAA CTAAGAA CACTGACTCTGAATGGTGCCTCAAAATACTGAA 960
Qy 961 TTTCTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Db 961 TTTCTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAAGTGTAGATCTGCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAAGTGTAGATCTGCT 1080
Qy 1081 TACAACTTATAGAAAGATTACCCAGTCTTTTCAGTCTGCCAAAAGCTTCAAGAAATTCAC 1140
Db 1081 TACAACTTATAGAAAGATTACCCAGTCTTTTCAGTCTGCCAAAAGCTTCAAGAAATTCAC 1140

Qy 1141 CTAAGACATATAGAAATCTACGAAATTTAAAGTTTGACACATTTTCCAGCAGTTGCTTAGCCTC 1200
Db 1141 CTAAGACATATAGAAATCTACGAAATTTAAAGTTTGACACATTTTCCAGCAGTTGCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTTATTTACCCCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTTATTTACCCCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTAAATAAGCTGGA CTTATCGTCCAAACCTCTGCTGCTTTTCCCTATAACT 1320
Db 1261 TTGCCATCCCTAAATAAGCTGGA CTTATCGTCCAAACCTCTGCTGCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGTTTAACTCACTTAAATTTAACAGAAATCATGCTTACAGAGCTTTGATA 1380
Db 1321 GGGTTACATGTTTAACTCACTTAAATTTAACAGAAATCATGCTTACAGAGCTTTGATA 1380
Qy 1381 TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGSCCTTATGCTTACCAAGTGTGT 1440
Db 1381 TCATCTGAAATCTTTCCAGAACTCAAGGTTATAGAAATGSCCTTATGCTTACCAAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGAGAAATGCTTAAGAATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTAAGAATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Qy 1501 AGCAGTATGGACGACCTTTTATAAGAAAGATCTGGAAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGGACGACCTTTTATAAGAAAGATCTGGAAATGTTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTTCTGCTTGA CTTTGGAGAAAGACTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTTCTGCTTGA CTTTGGAGAAAGACTGAAAGCCCTTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACTTCCAGAGGCCCTTCAACCCCTGTGAAACACCTGCTTGTAGTGGCTGGCTGATC 1680
Db 1621 TGTTCACTTCCAGAGGCCCTTCAACCCCTGTGAAACACCTGCTTGTAGTGGCTGGCTGATC 1680
Qy 1681 AGAATGGAGTGTGGACCATAGCAGTTCTGACACTTCTGCTTAATGCTTTGGTGAATTTCA 1740
Db 1681 AGAATGGAGTGTGGACCATAGCAGTTCTGACACTTCTGCTTAATGCTTTGGTGAATTTCA 1740
Qy 1741 ACAGTTTTTACAGATCCCTCTGTACATTTTCCCCCAATTTAAATTTGGGGTCAATCGCA 1800
Db 1741 ACAGTTTTTACAGATCCCTCTGTACATTTTCCCCCAATTTAAATTTGGGGTCAATCGCA 1800
Qy 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGTGGCTGGTGGTGGATGCGTTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGTGGCTGGTGGTGGATGCGTTCACT 1860
Qy 1861 TTTGGCAGCTTTGCAACGACATGGTCCCTGTTGGGAGAAATGGGGTTGGTGGCATGTCATT 1920
Db 1861 TTTGGCAGCTTTGCAACGACATGGTCCCTGTTGGGAGAAATGGGGTTGGTGGCATGTCATT 1920
Qy 1921 GGTTTTTTGTCCATTTTTTGTCTCAAGATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTTTGTCTCAAGATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCATTTTTTGTCTCTGCTGCTGCTTGGCATGCGCGCAGTTTCCCTCTG 2100
Db 2041 CTGAAAGTAAATCATTTTTTGTCTCTGCTGCTGCTTGGCATGCGCGCAGTTTCCCTCTG 2100
Qy 2101 CTGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCTGCTGCTTTGGCTTTTGGGAGGCC 2160
Db 2101 CTGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2160
Qy 2161 AGCACCATGGGCTACATGGTCTCATCTGCTCAATTTCCCTTGTCTCTCTCATGATG 2220
Db 2161 AGCACCATGGGCTACATGGTCTCATCTGCTCAATTTCCCTTGTCTCTCTCATGATG 2220

Qy	961	TTTCCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC	1020
Db	961	TTTCCTCACTTGA CAGGAACTGCCACCTGGAGAGTCTGACTTTTAACTGGAGCAAGATC	1020
Qy	1021	TCATCTCTTCTCTAAACGGTCTGCAATCAGTTA CTTAATCTCCAAGTGTCTAGATCTGCT	1080
Db	1021	TCATCTCTTCTCCCAAGGCGCTGTGATCAAGTTA CTTAATCTCCAAGTGTCTAGATTTGTCT	1080
Qy	1081	TACAACTATTAGAAAGATTTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAATTTGAC	1140
Db	1081	TACAACTACTCGAAGACTTACCAGTTTGT CAGGCTGCCAAAACCTTCAGAAATTTGAC	1140
Qy	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTTGCTTTAGCCTC	1200
Db	1141	CTGAGGCATAACGAGACTATAGAAATTAAGGCGCAGCACTTTTACGAGTTGTTTAACTCTC	1200
Qy	1201	CGATCGCTGAATTTTGGCTTTGGAAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT	1260
Db	1201	CGATCTCTGAACTTTAGACATGGAATTAATAATTTGCTATCAATTCACCCCAATTCGGTTTCTACG	1260
Qy	1261	TTGCGATCCCTTAATAAAGCTGGACCTATCGTCCAACCTCTCTGTCGTCTTTTCCCTATAACT	1320
Db	1261	TTGCGGCTCTPAATAAAGTTGGACCTATCATCCAATCTCTCTGTCGTCTTCCCTCTGACT	1320
Qy	1321	GGGTTACATGGTTTTAACTCACTTAAAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1321	GGGTTACATGGTTTTAACTCACTTAAAAATTAACAGGAACTTCAGAGCTTGATA	1380
Qy	1381	TCATCTGAAATCTTTCAGAACTCAAGGTTATAGAAATTTGCTTATGCTTACCAGTGTCTGT	1440
Db	1381	CCATCTGCAAACTTCCAGAGCTCAAGATTTATAGAAATGCCATCTGCTTACCAGTGTGT	1440
Qy	1441	GCAATTTGAGTGTGTGAGATGCTCTATAAGATTTTCTAATCAATGGAATAAAGGTGACAAC	1500
Db	1441	GCAATTTGGGGGTGTGAGATGCTTATAAATTTCTAAACCAATGGAATAAAGCAGCAGC	1500
Qy	1501	AGCAGTATGGA CGACCTTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1501	AACAGTGTGACGACCTTTCATAAGAAAGACGCTGGGTTATTTCAAGTTCAAGATGAGCGG	1560
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTAGTGCAG	1620
Db	1561	GACCTTGAAGATTTCTACTTTGACTTTTGAGGAAGACCTGAAAGCCCTTCACTCGTGCCAG	1620
Qy	1621	TGTTCACTTCCCGCAGCCCTTCAAACCTGTCAACACCTGCTTGATGGCTGGCTCATC	1680
Db	1621	TGCTCGCCTTTCCCGAGTCCCTTCAAGCCCTGTGAGCACCTATTGTGTAGCTGGCTGATC	1680
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTTGCGCACTTACTTTGTAATGCTTTGGTGACTTCA	1740
Db	1681	CGAATCGGGGTGTGGACCA CGG CAGTACTGGCGCTTTTCTGCAATGCTTTGGTGGCTTTG	1740
Qy	1741	ACAGTTTTCAGATCCCCTCTGTACA TTTCCGCCCATTAATACTGTTPAATTTGGGGTCAATCGCA	1800
Db	1741	ACCGTGTTTCAGAACTCTCTCTGTACATCTCTTCCAATAAAGCTGCTPAA TTGGGGTAAATCGCG	1800
Qy	1801	GCAGTGAACATGCTCA CGGAGTCTCCAGTGCCTGCTGGCTGTGTGTGATGCTGCTTCACT	1860
Db	1801	GTAGTGAACATTTCTCATGGGGGTCTCCAGTGTCTGTGTGGCTGGCTGGAATTCACCT	1860
Qy	1861	TTTGGCAGCTTTTGCA CGACATGTTGCCCTGTGGGAGAAATGGGGTTGGTTGCTCATGTCAATT	1920
Db	1861	TTTGGCGTTTTGTCTAGCAGCGTGGTGGTGGGAGACGGAAATCGGCTGCCAAATCGTT	1920
Qy	1921	GGTTTTTGTCCATTTTGTCTCAGAAATCAATCTGTTTTCTGTCTTACTCTGGCAGCCCTG	1980
Db	1921	GGCTTCTGTCCATTTTGTCTCCGAATCGTCGATCTTCTCTGTCTCACTCTGGCAGCGCTG	1980
Qy	1981	GAGCGTGGGTCTCTGTCAATATTTCTGCAAAAATTTTGAACGGAAGCTCCATTTTCTAGC	2040
Db	1981	GAA CGAGGTTTTCTGTCAAGTGTCTTTCGAAGTTTGAAGTGAAGCTCCCTTTTTTATGC	2040
Qy	2041	CTGAAAGTAATCATTTTGTCTGTGCCCTTGTCTGCCCTTGACATGCGCGGAGTTTCCCTGT	2100

[illegible]

Adams, M.D. and Cargill, M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. FEATURES source Location/Qualifiers 1..2593 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>2593 /gene="GPR49" /locus_tag="HCW4322"									
ORIGIN									
Query Match 71.9%; Score 1958; DB 9; Length 2593; Best Local Similarity 75.7%; Pred. No. 0; Matches 1964; Conservative 0; Mismatches 629; Indels 0; Gaps 0;									
QY	132	CAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTGCCTTCCAACT	191						
DB	1	CAGGATGTTGCTCAGGGTNNNNNNNTCCGACNNNGGCTCTCGAGCTGCCTTCCAACT	60						
QY	192	GAGCGCTTTCACCTCTACTAGACCTCAGTATGAACAACATCAGTCAGCTGCTCCGAA	251						
DB	61	GAGCGCTTTCACCTCTACTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	120						
QY	252	TCCCTGCGCCAGTCTCCGCTCTCTGGAGGAGTTACGTCTTGGCGGAAACGCTCGACATA	311						
DB	121	NN	180						
QY	312	CATTCCTCCAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTATGCTGCAGATAA	371						
DB	181	NN	240						
QY	372	TCAGTAAGACAGTACCACAGAGCTCTGCAGAAATTTGGAGAGCTTCAATCCCTGGC	431						
DB	241	NN	300						
QY	432	TCGTGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGCGTCGATTCCT	491						
DB	301	TCGTGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGCGTCGATTCCT	360						
QY	492	GAGGCACCTGTGGCTGGATGACATGCGTTTAAACAGAAATCCCGCTCCAGGCTTTTAGAAG	551						
DB	361	GAGGCACCTGTGGCTGGATGACATGCGTTTAAACAGAAATCCCGCTCCAGGCTTTTAGAAG	420						
QY	552	TTTATCGGCATTGCGAAGCATGACCTTGGCCCTGAAACAAATAACACATACCGAGCTA	611						
DB	421	TTTATCGGCATTGCGAAGCATGACCTTGGCCCTGAAACAAATAACACATACCGAGCTA	480						
QY	612	TGCGCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACATAGAAATCCACTC	671						
DB	481	TGCGCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACATAGAAATCCACTC	540						
QY	672	CCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTCTAGAGACTTTTAGATTTAAATACAA	731						
DB	541	CCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTCTAGAGACTTTTAGATTTAAATACAA	600						
QY	732	TAACTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAAGAACTAGGATT	791						
DB	601	NN	660						
QY	792	TCATAGCAATATCAGTTCGATACCTCAGAAAGCATTTGTAGGCAACCTCTCTTAT	851						
DB	661	TCATAGCAATATCAGTTCGATACCTCAGAAAGCATTTGTAGGCAACCTCTCTTAT	720						
QY	852	TACAAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTT	911						
DB	721	TACAAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTT	780						
QY	912	ACCTGAACTAAGAACACTGACTCTCGAATGGTGGCTCCACAAATAACTGAATTTCTGATTT	971						

DB	781	ACCTGAACTNN	840						
QY	972	AACTGGAACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCATCTCTTCC	1031						
DB	841	AACTGGAACTNN	900						
QY	1032	TCAAACCGTCTGCAATCAGTTACCTAAATCTCCAAGTGTAGATCTGTCTTACAACTATT	1091						
DB	901	NN	960						
QY	1092	AGAAGATTATCCACAGTTTTCAGTCTGCCAAAGCTTCCAGAAATTTGACTAGACATAA	1151						
DB	961	AGAAGATTATCCACAGTTTTCAGTCTGCCAAAGCTTCCAGAAATTTGACTAGACATAA	1020						
QY	1152	TGAAATCTACGAAATTTAAAGTTGACACATTTCCAGCAGTTGCTTAGGCTCCGCTCGTGAA	1211						
DB	1021	TGAAATCTNN	1080						
QY	1212	TTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTTCCACTTGGCATCCCT	1271						
DB	1081	NN	1140						
QY	1272	AATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATAACTGGGTTCATGG	1331						
DB	1141	NN	1200						
QY	1332	TTTAACTCACTTAAATTTAAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTGAAA	1391						
DB	1201	TTTAACTCACTTAAATTTAAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTGAAA	1260						
QY	1392	CTTTCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACCAAGTGTGTGCAATTTGGAGT	1451						
DB	1261	CTTTCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACCAAGTGTGTGCAATTTGGAGT	1320						
QY	1452	GTGTCAGATGCTATAAGATTTCTAATCAATGGAATAAAGTGAACAACAGCATATGGA	1511						
DB	1321	GTGTCAGATGCTATAAGATTTCTAATCAATGGAATAAAGTGAACAACAGCATATGGA	1380						
QY	1512	CGACCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACCTGCACTTGAAGA	1571						
DB	1381	CGNCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACCTGCACTTGAAGA	1440						
QY	1572	TTTCTGCTGTGCTTTGAGGAGACCTGAAAGCCCTTCAATCAGTGCAGTGTTCACCTTC	1631						
DB	1441	TTTCTGCTGTGCTTTGAGGAGACCTGAAAGCCCTTCAATCAGTGCAGTGTTCACCTTC	1500						
QY	1632	CCAGGCCCTTCAACCCCTGTGAACACTGCTTCAATGGCTGGCTGATCAGAAATTTGGAGT	1691						
DB	1501	CCAGNN	1560						
QY	1692	GTGGACCATAGCAGTCTGGCACTTACTTGTAACTGTTGGTGACTTCAACAGTTTTCAG	1751						
DB	1561	NN	1620						
QY	1752	ATFCCCTCTGTAATTTCCCACTTAAACTGTTAAATGGGGTTCATCGCAGCATGAACAT	1811						
DB	1621	NN	1680						
QY	1812	GCTCAGGGAGTCTCAGTGCCTGCTGGTGGTGGATGCGTTCACCTTTTGGCAGCTT	1871						
DB	1681	GCTCAGGGAGNNNNNAGTGGCGTGGTGGATGCGTTCACCTTTTGGAGAGCTT	1740						
QY	1872	TGCAGACATGTTGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1931						
DB	1741	TGCAGACATGTTGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1800						
QY	1932	CATTTTTCCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTGGAGCGGGTT	1991						
DB	1801	CATTTTTCCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTGGAGCGGGTT	1860						
QY	1992	CTCTGTGAAATTTCTGCAAAATTTGAAACGAAAGCTCCTATTTCTAGCCTGAAAGTAT	2051						

QY 1718 CTTGTAATGCTTTGGTGAATCAACAGATTTTCAGATCCCTCTGTACATTTCCCCCATTA 1777
 Db 721 CTTGTAATGCTTTGGTGAATCAACAGATTTTCAGATCCCTCTGTACATTTCCCCCATTA 780
 QY 1778 AACTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGC 1837
 Db 781 AACTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGC 840
 QY 1838 TGGCTGGTGTGATGCGTTCACTTTTGGCAGCTTTTGGCAGCATGCTGCTGGTGGGAGA 1897
 Db 841 TGGCTGGTGTGATGCGTTCACTTTTGGCAGCTTTTGGCAGCATGCTGCTGGTGGGAGA 900
 QY 1898 ATGGGGTGGTGTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1957
 Db 901 ATGGGGTGGTGTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 1958 TCCTGCTTACTCTGCGAGCCCTGGAGCGTGGGTTCTCTGTAATTAATTTCTGCAAAATTTG 2017
 Db 961 TCCYGCYTACTCYKSMGCCCTGGAGCGTGGGTTCTCTGTAATTAATTTCTGCAAAATTTG 1020
 QY 2018 AAACGAAAGCTCCATTTTCTAGCTGGAAGTAATCAATTTGCTGCTGCTGCTGCTGCTGCT 2077
 Db 1021 AAACGAAAGCTCCATTTTCTAGCTGGAAGTAATCAATTTGCTGCTGCTGCTGCTGCTGCT 1079
 QY 2078 TGACCATGCGCGCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 Db 1080 GACCHGCGCCGTTTCCCTCCCGG 1102

RESULT 5

AL541959
 DEFINITION AL541959 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB007YC19
 5-PRIME, mRNA sequence.
 ACCESSION AL541959
 VERSION AL541959.3 GI:45717535
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1060)
 Li,W.B., Gruber,C., Jesssee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:30546637.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 7021.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?as=CS0DB007AB10QPI&c=7021.f.

FEATURES

Location/Qualifiers
 1..1060
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB007YC19"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 35.9%; Score 979.2; DB 1; Length 1060;
 Best Local Similarity 98.3%; Pred.No. 5.4e-273;
 Matches 993; Conservative 6; Mismatches 9; Indels 2; Gaps 1;
 QY 683 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 742
 Db 1 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 60
 QY 743 AATTTCCCACTGCAATTAAGGACATCTCTCAACCTTTAAAGAACTAGATTTTCATAGCAACA 802
 Db 61 AATTTCCCACTGCAATTAAGGACATCTCTCAACCTTTAAAGAACTAGATTTTCATAGCAACA 120
 QY 803 ATATCAGGTCCTACCTGAGAAAGCATTTGTAGGCAACCTCTCTTATTACAAATACATTT 862
 Db 121 ATATCAGGTCCTACCTGAGAAAGCATTTGTAGGCAACCTCTCTTATTACAAATACATTT 180
 QY 863 TCTATGACAAATCCCAATCTCAATTTGTGGAGATCTGCTTTTCAACATTTTACCTGAACTAA 922
 Db 181 TCTATGACAAATCCCAATCTCAATTTGTGGAGATCTGCTTTTCAACATTTTACCTGAACTAA 240
 QY 923 GAAACATGACTCTGAAATGGTGCCTCAAAATTAACCTGAATTTCTGATTTAACTGGAATG 982
 Db 241 GAAACATGACTCTGAAATGGTGCCTCAAAATTAACCTGAATTTCTGATTTAACTGGAATG 300
 QY 983 CAACCTGGAGAGTCTGACTTTAACTGAGCAGCATCTCATCTCTTCTCAACCGCTCT 1042
 Db 301 CAACCTGGAGAGTCTGACTTTAACTGAGCAGCATCTCATCTCTTCTCAACCGCTCT 360
 QY 1043 GCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCTTTACAACTATTAGAAAGATTTAC 1102
 Db 361 GCAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCTTTACAACTATTAGAAAGATTTAC 420
 QY 1103 CCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGACCTAAGACATATGAAATCTACG 1162
 Db 421 CCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGACCTAAGACATATGAAATCTACG 480
 QY 1163 AAATTAAGTTGACACATTTCCAGCAGTGTGCTTAGCTCCGATCGTGATTTGGCTTGA 1222
 Db 481 AAATTAAGTTGACACATTTCCAGCAGTGTGCTTAGCTCCGATCGTGATTTGGCTTGA 540
 QY 1223 ACAAATTTGCTATTATTCAACCCCAATGATTTTCCACATTTGCTCATCTCCATTAAGAGCTGG 1282
 Db 541 ACAAATTTGCTATTATTCAACCCCAATGATTTTCCACATTTGCTCATCTCCATTAAGAGCTGG 600
 QY 1283 ACCTATCGTCCAACTCTGCTGCTTTTCTATAACTGGGTTACATGTTTAACTCACT 1342
 Db 601 ACCTATCGTCCAACTCTGCTGCTTTTCTATAACTGGGTTACATGTTTAACTCACT 660
 QY 1343 TAAATTAAGGAAATCATGCTTACAGACTTGATATCATCTGAAAACCTTTCCAGAAC 1402
 Db 661 TAAATTAAGGAAATCATGCTTACAGACTTGATATCATCTGAAAACCTTTCCAGAAC 720
 QY 1403 TCAAGGTTTATAGAATGCTTATGCTTACAGTGTGTCATTTGGAGTGTGAGAAATG 1462
 Db 721 TCAAGGTTTATAGAATGCTTATGCTTACAGTGTGTCATTTGGAGTGTGAGAAATG 780
 QY 1463 CCTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAGCAGTATGACGACCTTCATA 1522
 Db 781 CCTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAGCAGTATGACGACCTTCATA 840
 QY 1523 AGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACCTGAGCTTGAAGATTTCTGCTTG 1582
 Db 841 AGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACCTGAGCTTGAAGATTTCTGCTTG 900
 QY 1583 ACTTTGGAGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCGAGCCCT 1642
 Db 901 ACTTTGGAGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCGAGCCCT 958
 QY 1643 TCAAAACCTGTGAACACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
 Db 959 TCAAAACCTGTGAACACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008


```
Db 750 CTGCACTCCCTGAGGCACCTGTGGCTAGATGACAAATGCTCTCACAGACGTCCTGTCAC 809
Qy 541 GCTTTTGAAGTTATCGGATTCGAAGCATGACCTTGGCCCTGAAACAAATACACAC 600
Db 810 GCTTTTGAAGTTATCGGATTCGAAGCATGACCTTGGCCCTGAAACAAATACACAC 869
Qy 601 ATACACAGATGTCCTTTGGAACCTCTCCAGCTTGGTAGTTCTACATCTCCATACAT 660
Db 870 ATAGCAGATGTCCTTTGGAACCTCTCCAGCTTGGTAGTTCTACATCTCCATACAT 929
Qy 661 AGAATCCACTCCCTGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 930 AGAATCCACTCCCTGGAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 989
Qy 721 TTAATATACAAATCTGATGATTTCCCACTGCAATTAGGACACTCTCCAACTTAA 780
Db 990 TTAATATTAATAACCTTTGATGAATTTCCCACTGCAATCAAGACACTCTCCAACTTAA 1049
Qy 781 GAACTAGGATTTATAGCAACCAATATCAGTGTGATACCTGAGAAAGCAATTTAGGCAAC 840
Db 1050 GAACTAGGATTTATAGCAACCAATATCAGTGTGATACCTGAGAAAGCAATTTAGGCAAC 1109
Qy 841 CTTCTCTTATTAACATATTTATAGCAATTTCCATCCATCCCAATTTGTTGGAGATCTGCT 900
Db 1110 CTTCTCTTATTAACATATTTATAGCAATTTCCATCCATCCCAATTTGTTGGAGATCTGCT 1169
Qy 901 TTTCAACATTTACTGGAATAGAACACTGACTCTGAATGGTGCCTTCAAAATACGAA 960
Db 1170 TTTCAACATTTACTGGAATAGAACACTGACTCTGAATGGTGCCTTCAAAATACGAA 1229
Qy 961 TTTCTGATTTTAACTGGAACTGCAACCTGGAGACTGACTTAACTGGAGCAGATC 1020
Db 1230 TTTCTGATTTTAACTGGAACTGCAACCTGGAGACTGACTTAACTGGAGCAGATC 1289
Qy 1021 TCATCTCTTCTTCAAACTGTCGCAATCAGTTACCTTAATCTCAAGTGTAGATCTGCT 1080
Db 1290 TCATCTCTTCTTCAAACTGTCGCAATCAGTTACCTTAATCTCAAGTGTAGATCTGCT 1349
Qy 1081 TACAACCTATTAGAGATTTACCCAG 1106
Db 1350 GTAAGGCAATGATGTTGTATATAACAG 1375

RESULT 7
LOCUS AK085901
DEFINITION Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830026M09 product:CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens], full insert sequence.
ACCESSION AK085901
VERSION AK085901.1 GI:26103061
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
```

```
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Hara,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3583)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,M., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saichon,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tanaka,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.jp, URL:http://genome.gsc.riken.jp/, Tel.81-45-503-9222, Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
1..3583
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D830026M09"
/db_xref="taxon:10090"
/clone="D830026M09"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="16 days neonate"
95..3002
/note="CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens] (SPR[Q96K69], evidence: FASTY, 85.8%ID, 92.1%length, match=637)
putative"
3563..3568
/note="putative"
3583
/note="putative"
misc_feature
polyA_signal
polyA_site
ORIGIN
```


D _b	2208	CGCTGGCGCTCGGTGGGAGAGTATGGCGCCTCCCCACCTCGCTGCGCCTACGCCCCACCGC	2267
Q _y	2156	AGCCACG-----CACCATGGGCTACATGGTCGCTCTCATCTTGTGCTCAATTCCTTTGTC	2208
D _b	2268	AGGCGCGCGCGCGCTGCGCTGGGCTTCGCTGTAGACCCCTGGTGATGATGAACCTCGCTCTGC	2327
Q _y	2209	TTCTCATGATGACCATTTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTG	2268
D _b	2328	TTCTGTGTGTGGCGCGCGCTACATCAAGCTCTACTGTGACCTGCGCACGGGGTGACTTT	2387
Q _y	2269	GAGAAATTTGGGACCTGCTCTATGGTAAAAACATATGCCCTGTGTGCTTCTTCACCAACTGC	2328
D _b	2388	GAGSCCGTGGGACTCGCGCATGGTGGCGCACGTGSCCTGCTCATCTTTGCAGATGCG	2447
Q _y	2329	ATCGTAAACGTCCCTGTGGCTTTCTTGTCCTCTCTCTTTAATAAACCTTACATTTATC	2388
D _b	2448	CTCCTCTACTGCCCCGGCGCTTCCTCAGCTTTTGGCCTCTAATGCTGGGCGCTTCTCCCTGTC	2507
Q _y	2389	AGTCTCGAAGTAAATTAAGTTTATCTTCTCGTGGTAGTCCACCTTCCTCGCATGCTCAAT	2448
D _b	2508	ACCCCGAGCGCTGTCAAGTCAGTCCCTTCGTGGTGTGCTCTGTGCTGCTGCTGCCTCAAC	2567
Q _y	2449	CCGCTTCTTACATCTTTGTTCAATCTCTCACTTTAAGGAGGATCTGNGAGCCT	2501
D _b	2568	CCACTGCTTACCTGCTCTTAAACCCCTCACTTCGGGATGACCTTCGGCGGCT	2620

RESULT	8
CK449409	
LOCUS	CK449409 940 bp mRNA linear EST 12-JAN-2004
DEFINITION	892976 MARC 4P1G Sus scrofa cDNA 5' , mRNA sequence.

ORGANISM	SUS scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
REFERENCE	1 (bases 1 to 940)
AUTHORS	Smith, T.P.L., Preking, B.A., Ford, J.J., Vallet, J.L., Wise, T.A., Nonneman, D.J., Wray, J.E. and Keele, J.W.
TITLE	Porcine EST collection using a normalized library constructed from embryos representing early developmental stages

ORIGIN

Query Match	27.4%	Score 745.2;	DB 7;	Length 940;
Best Local Similarity	86.6%	Pred. No. 7.6e-205;		
Matches 876; Conservative	0;	Mismatches 63;	Indels 73;	Gaps 2;

RESULT	9
BGI96708	
LOCUS	BGI96708 800 bp mRNA linear EST 21-APR-2001
DEFINITION	RST15934 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION BG196708
VERSION BG196708.1 GI:13718395
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Harrington,J.J., Sherr,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,K., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc. Ave, Cleveland, OH 44115, USA
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 481.
FEATURES
source
1..800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
ORIGIN
Query Match 27.1%; Score 738.6; DB 4; Length 800;
Best Local Similarity 98.9%; Pred. No. 5.9e-203;
Matches 796; Conservative 0; Mismatches 4; Indels 5; Gaps 5;
Qy 1210 AATTGGCTTGGACAAAATTGCTATTATTCACCCCAATGCAATTTCCACTTTGCCATCC 1269
Db 1 AATTGGCTTGGACAAAATTGCTATTATTCACCCCAATGCAATTTCCACTTTGCCATCC 60
Qy 1270 CTAATAAAGCTGGACCTATCGTCCAACTCCTGCTGCTTTTCCCTATAACTGGGTTACAT 1329
Db 61 CTAATAAAGCTGGACCTATCGTCCAACTCCTGCTGCTTTTCCCTATAACTGGGTTACAT 120
Qy 1330 GGTTTAACTCACTTAAAAATTAAACAGGAATCATGCTTTACAGAGCTTCATATCATCTCAA 1389
Db 121 GGTTTAACTCACTTAAAAATTAAACAGGAATCATGCTTTACAGAGCTTCATATCATCTCAA 180
Qy 1390 AACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGCTGCTTTGGA 1449
Db 181 AACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGCTGCTTTGGA 240
Qy 1450 GTGTGTGAGATGCCCTATAAGATTCTTAATCAATGGAATAAGGTGACAAACAGCAGTATG 1509
Db 241 GTGTGTGAGATGCCCTATAAGATTCTTAATCAATGGAATAAGGTGACAAACAGCAGTATG 300
Qy 1510 GACGACCTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGTCGACCTGAA 1569
Db 301 GACGACCTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGTCGACCTGAA 360
Qy 1570 GATTTCCTGCTGACTTTGAGGAAAGCCTGAAGCCCTTCATTCAGTCGAGTGTTCACCT 1629
Db 361 GATTTCCTGCTGACTTTGAGGAAAGCCTGAAGCCCTTCATTCAGTCGAGTGTTCACCT 420
Qy 1630 TCCCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGCTTGATGGCTGCTGATCAGAATTGGA 1689

Db 421 TCCCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGCTTGATGGCTGCTGATCAGAATTGGA 480
Qy 1690 GTGTGGACCATAGCAGTTCTGGCACTTACTTGTATGCTTTGGTACTTCAACAGTTTTC 1749
Db 481 GTGTGGACCATAGCAGTTCTGGCACTTACTTGTATGCTTTGGTACTTCAACAGTTTTC 540
Qy 1750 AGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGTTCATCGCAGCAGTGAAC 1809
Db 541 AGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGTTCATCGCAGCAGTGAAC 600
Qy 1810 ATGCTCAAGGAGTCTCCAGTCCGCTGCTGCTGGTGGTGGATGCGTTCATTTTGGCAGC 1869
Db 601 ATGCTCAAGGAGTCTCCAGTCCGCTGCTGCTGGTGGTGGATGCGTTCATTTTGGCAGC 660
Qy 1870 TTTCACAGCATGCTGCTGGTGGAGAAATGGGGTGGTTCGCCATGTCATTTGTTTGTG 1929
Db 661 TTTCACAGCATGCTGCTGGTGGAGAAATGGGGTGGTTCGCCATGTCATTTGTTTGTG 720
Qy 1930 TCCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTGGAGCGTGGG 1989
Db 721 -CCATTTTGTCTCAG-ATCATCTGTTTTCCTG-TTACTCTGGC-GGCTCGAGCGTGGG 776
Qy 1990 TTCTCTGTGAATATTCTGCAAAAT 2014
Db 777 TTCTCTGTG-AATATTCTGCAAAAT 800
RESULT 10
BM795259
LOCUS K-EST0077025 S21SNU520 Homo sapiens cDNA clone S21SNU520-31-C01 5',
DEFINITION mRNA sequence.
ACCESSION BM795259
VERSION BM795259.1 GI:19143491
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 671)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: C column: 01
High quality sequence stop: 671.
FEATURES
Location/Qualifiers
1..671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="S21SNU520-31-C01"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNU520"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transposition of competent cells E. coli Top10⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 24.6%; Score 669.4; DB 4; Length 671;
Best Local Similarity 99.9%; Pred. No. 8e-183;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1775 TTAACCTGTTAAATGGGTCATCGCAGCATGCAATGCTCACGGAGTCTCCAGTCCG 1834
Db 1 TTAACCTGTTAAATGGGTCATCGCAGCATGCAATGCTCACGGAGTCTCCAGTCCG 60
QY 1835 TGCTGGCTGGTGGATGGTTCACCTTTTGGCAGCTTTTGCACGACATGCTGCTGGTGG 1894
Db 61 TGCTGGCTGGTGGATGGTTCACCTTTTGGCAGCTTTTGCACGACATGCTGCTGGTGG 120
QY 1895 AGAATGGGTTGGTGGCAGTCAATGTTTTTGGTTCATTTTGGTTCAGAAATCATCTG 1954
Db 121 AGAATGGGTTGGTGGCAGTCAATGTTTTTGGTTCATTTTGGTTCAGAAATCATCTG 180
QY 1955 TTTTCTCTGTTACTTGGCAGCCCTGGAGCGTGGTCTCTGTGAATATTTGCAAAAT 2014
Db 181 TTTTCTCTGTTACTTGGCAGCCCTGGAGCGTGGTCTCTGTGAATATTTGCAAAAT 240
QY 2015 TTGAACGAAAGCTCCATTTTCTAGCTGAAAGTAAATCATTTTGTCTGTGCTGCTGG 2074
Db 241 TTGAACGAAAGCTCCATTTTCTAGCTGAAAGTAAATCATTTTGTCTGTGCTGCTGG 300
QY 2075 CTTGACCATGCGCGAGTTCCTCTGCTGGTGGGAGCAAGTATGCGGCTCCCTCTCT 2134
Db 301 CTTGACCATGCGCGAGTTCCTCTGCTGGTGGGAGCAAGTATGCGGCTCCCTCTCT 360
QY 2135 GCCTGCTTGGCTTTTGGGAGCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2194
Db 361 GCCTGCTTGGCTTTTGGGAGCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 2195 TCAATTCCTTGGCTTCTCATGATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2254
Db 421 TCAATTCCTTGGCTTCTCATGATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 2255 ACAAGGAGACTGAGAAATATTTGGGACTGCTCTATGTAATAACACATTTGCTGCTG 2314
Db 481 ACAAGGAGACTGAGAAATATTTGGGACTGCTCTATGTAATAACACATTTGCTGCTG 540
QY 2315 TCTTCAACACTGCTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2374
Db 541 TCTTCAACACTGCTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 2375 ACCTTACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2434
Db 601 ACCTTACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 2435 CTGCAATGCTC 2445
Db 661 CTGCAATGCTC 671

RESULT 11

BC038795
LOCUS
DEFINITION
Homo sapiens leucine-rich repeat-containing G protein-coupled receptor 6, mRNA (cDNA clone IMAGE:5220507), with apparent retained intron.
ACCESSION
BC038795
VERSION
BC038795.1 GI:24433474

KEYWORDS

SOURCE
ORGANISM

REFERENCE
AUTHORS

HTC.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 3509)

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Kyselynski, M.I., Skalek, U., Smalish, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3509)

Strausberg, R.

Direct Submission

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC)

Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@hgrl.nih.gov

Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAC Plate: 64 Row: n Column: 17

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis

This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers

1..3509

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5220507"

/tissue_type="Pancreas, Spleen, adult pooled"

/clone_lib="NIH_MGC_120"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 24.3%; Score 661.4; DB 3; Length 3509;
Best Local Similarity 57.1%; Pred. No. 3.4e-180;
Matches 1388; Conservative 0; Mismatches 911; Indels 130; Gaps 5;

QY 200 TCACCTCTACTAGACCTCAGTAGTGAACAACATCAGTCAGTCGTCCCGAATCCCTGCG 259
DB 87 TCAITCTGCGCGGACCTCAGATGAACAACCTCAGAGCTTCAGCTGGCTCTTCC 146

QY 260 CCAGTCTCCGCTCTCTGAGGAGTTACGTCTTGCGGAAACGCTCTGACATACATPCCCA 319
DB 147 ACCACTCGCTCTCTTGAGGAGCTGCGTCTCTCTGGGAACCATCTCTCACACATCCAG 206

QY 320 AGGAGGCAATCCTGCGCTTTACAGTCTTAAGTCTTATGCTGAGAGTAATCAGCTAA 379
DB 207 GACAGCAATCTCTGCTCTACAGCTGAAATCTCTGATGCTGAGAAACATCAGCTGG 266

QY 380 GACAGTACCCACAGAAGCTCTGCAGAATTTGCGAAGCTTCAATCCCTGCGTCTGAGTG 439
DB 267 GAGGATCCCGCAGAGGCGCTGTGGAGCTGCCGAGCTCGAGTCTGCGCTAGATG 326

QY 440 CTAACCAATCAGCTATGTGCCCCCAAGCTGTTCAGTGGCTGCAATCCCTGAGGCACC 499
DB 327 CCAACTCTCTCCCTGCTCCGAGAGAGAGCTTTGAGGGCTGTCTCTCCCTCCGCCACC 386

QY 500 TGTGCTGGATGACATGGTGTAAACAGAAATCCCGCTCAGGCTTTTGAAGTTTATCGG 559
DB 387 TCTGCTGGAGCAATGACATCAGCGAGATCCCTGTGAGGGCCCTCAACACCTCCCTG 446

QY 560 CATTCGAAGCATGACCTTGGCCCTGAAACAAATAACACATACAGACTATGCTTTG 619
DB 447 CCCTGAGGCATGACCTTGGCCCTCAACCGCATCAGCACATCCCGACTACGGTTCC 506

QY 620 GAAACCTCTCAGCTTGTGTAGTTCTACATCTCCATCAACAATAGAAATCCATCCCTGGAA 679
DB 507 AGAATCTCACAGCCTTGTGTGTGCTGATTTGTCATAACAAACCGCATCAGCATCTGGGA 566

QY 680 AGAATGCTTTGATGGCTTCCACAGCTTAGAGACTTTAGATTTAAATTAACATACCTTG 739
DB 567 CCCACAGCTTCGAGGGGTGCACAACTCTGGAGACATAGACTGAATTTATAACAGCTGC 626

QY 740 ATGAATTCACCACTGCAATTAGGACACTCTCCAACTTTAAAGAACTAGGATTTATAGCA 799
DB 627 AGGAGTTCCTGTGCGCATCCGACCTTGGGAGACTCGAGNACTGGGGTTCCATAACA 686

QY 800 ACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAACCTTCTCTTATTAACATAC 859
DB 687 ACAACATCAAGGCCATCCAGAAAGGCTTTCATGGGAAACCTCTGCTACAGACGATAC 746

QY 860 ATTTCTATGACATCCGATCCAAATTTGTTGGAGATCTGCTTTTCAAATTTACTGAAAC 919
DB 747 ACCTTTATGATAACCAATCCAGTTTGTGGGAAGATCGGCATTCAGTACCTGCGCTAAAC 806

QY 920 TAAGAACTACTGACTCTGAATGTGCTCACAATAAATCTGAATTTCTGATTTAACTGGAA 979
DB 807 TCCACAGACTATCTCTGAATGTGCGCATGGACATCCAGAGTTTCCAGATCTCAAGGCA 866

QY 980 CTGCAAACTGAGAGTCTGACTTTTAACTGGAGCAAGATCTCATCTCTTCTCTAAACCG 1039
DB 867 CCACAGCTGGAGATCTGACCTGACCCGCGCAGGCATCCGGCTGCTCCCATCGGGA 926

QY 1040 TCTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGCTTACAACTTATAGAGATT 1099
DB 927 TGTGCAACAGCTGCCAGGCTCCGAGTCTCGAACTGTCTCACAATCAAAATTTAGGAGC 986

QY 1100 TACCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTT----- 1137
DB 987 TGCCCAAGCTGACAGAGTGTGAGAAATTTGGAGAAATGAGAGTGGGGGCTCTACTTTC 1046

QY 1138 ----- 1137
DB 1047 TTCTCCAGTGTCTCGGGGGGCAATGGAGCGGAGCCAGGCTGTGAGCCTGCGGCTCATCC 1106

QY 1138 -----GACCTAAGACATAATGAAATCTAGAAATTTAAAGTTGAC 1176
DB 1107 AGCTCTCTTCTGCTGCCCTAGCGGCTTCCAACACACCGCATCTGGAAATTTGGAGCTGAC 1166

QY 1177 ACTTTCCAGCAGTGTGCTTAGCTCCGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATT 1236
DB 1167 ACCTTCAGCAGCTGAGCTCCCTGCAAGCCTGGATCTTAGCTGGAACGCCATCCGGTCC 1226

QY 1237 ATTCAACCCCAATGATTTTCCACTTTGGCATCCCTAATAAAGCTGGAGCTATCGTCCAAAC 1296
DB 1227 ATCCACCTCGAGGCTTCTCCACCTGCACTCCCTGGTCAAGCTGGACCTGACAGACAAC 1286

QY 1297 CTCTGTCTGCTTTTCTTATTAACCTGGGTTACATGTTTTAACTCACTTAAAAATTAACAGGA 1356
DB 1287 CAGCTGACCAACATGCCCCCTGGCTGGACTTGGGGCTTGTGATCTGAGCTCAAGGG 1346

QY 1357 AATCATGCTTTCAGAGCTTGTATATCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAA 1416
DB 1347 AACCTTGTCTCTCCAGGCTTCTCCAAGGACAGTTTCCCAAAACCTGAGGATCCTGGAG 1406

QY 1417 ATGCTTATGCTTACAGTGTCTGTGCAATTTGGAGTGTGAGATGCTATAGATTTCT 1476
DB 1407 GTGCTTATGCTTACAGTGTCTGCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCT 1466

QY 1477 AATCAATGGAATAAAGGTGACAAACAGCATATGACGACCTTTCATAAGAAAGA----- 1529
DB 1467 GGGCAGTGGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTG 1526

QY 1530 -----TGCTGGAATGTTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGTCT 1581
DB 1527 GGCCTCTCTGCCAGACAAGCAGAGAACCACTATGACAGGACCTGGATGAGTCTCCAGCTG 1586

QY 1582 GACTTTGAGGAAGACCTGAAAGCCCTTCAITTCAGTGCAGTGTTCACCTTCCCAGGCCCC 1641
DB 1587 GAGATGGAG---GACTCAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCC 1643

QY 1642 TTCAAAACCTCTGAACACCTCTGTTGATGGCTGGCTGATCAGAAATGAGAGTGTGAACCATTA 1701
DB 1644 TTCAAGCCCTGTGAGTACTCTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATC 1703

QY 1702 GCAGTTCGCGACTTACTTTGTAATGCTTTGTTGATCTTCAACAGTTTTTC---AGATCCCT 1758
DB 1704 GTGTTGCTCTCGTGTCTCTGCAATGGACTGGTGTGCTGACCGTGTTCGCTGGCGGCT 1763

QY 1759 CTGTACATTTCCCCCATTAACACTGTTAATTTGGGCTCATCGCAGCAGTGAACATGCTCAG 1818
DB 1764 GCGCCCTGCGCCCGTCAAGTTTGTGTAGGTGCGATTTGAGGCGCCCAACACCTTGACT 1823

QY 1819 GGAGTCTCCAGTGCCTGCTGCTGGTGTGATGCTTCACTTTTGGCAGCTTTTGACGA 1878
DB 1824 GGCAATTTCTGTGGCTTCTAGCCTCAGTCAATGCCCTGACCTTTGGTCAGTCTCTGAG 1883

QY 1879 CATGTGCTGGTGGGAGAAATGGGTTGGTTGCCATGTCAATGGTTTTTTTGTCCATTTTT 1938
DB 1884 TACGAGCCCGCTGGAGACGGGCTAGGCTGCCGGGCACTGGCTTCTCTGSCAGTACTT 1943

QY 1939 GCTTCAGNATCATCTGTTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGTCTCTCTG 1998
DB 1944 GGGTCGAGGATCGGTGCTGCTCACTCTGCGCGCAGTGTGAGTGTGAGGCTCTCCGTC 2003

QY 1999 AAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAAATCAATTTG 2058
DB 2004 TCCTGTGTCGGGCTATGGGAAGTCCCTCTCCCTGGCAGGGTTCGAGCAGGGTCTCTA 2063

QY 2059 CTCTGTGCCCCCTGCTGGCTTGAACCATGGCCGCAATTTCCCTGCTGGTGGGAGCAAGTAT 2118
DB 2064 GGTGCTGGCACTGGCAGGGCTGGCGCGCCGCACTGCCCCCTGGCTCAGTGGGAGAAATAC 2123

QY 2119 GGGCTTCCCTCTCTGCTGCTT-----TGCTTTTGGGAGCCCGACCATATGGC 2172
DB 2124 GGGGCTTCCCACTCTGCTTCCCTTACGCGCAACCTGAGGGTTCAGCCAGCAGCCCTGGC 2183

QY 2173 TACATGGTGGCTCTCATCTTGTCTCAATCCCTTGTCTTCTCATGATGACCAATGCTCTAC 2232
Db 2184 TTACCGTGGCCCTGGTGTATGATGAATCTTCTTGTCTTCTGGTGGCGGTCCTAC 2243
QY 2233 ACCAAGCTCTACTGCAATTTTGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATG 2292
Db 2244 ATCAAACTGTACTGTGACCTCGCGGGGCGACTTTTGAGCGCGTGTGGACTGGCGCATG 2303
QY 2293 GTAAACACATTTGCCCTGTGTCTTACCAAACTGCATCTTAAACTGCGCTGTGGCTTTC 2352
Db 2304 GTGAGGACATGGCCCTGGCTCATCTTCGAGAGGGGCTCTCTACTGTCCCGTGGCTTC 2363
QY 2353 TTGTCTTCTCTCTTTAAATAAACCCTTACATTTATCAGTCTTGAAGTAATTAAGTTTATC 2412
Db 2364 CTCAGCTTCCCTCCATGCTGGGCTCTTCCCTGTGACGCGCGGCGCTCAAGTCTGTC 2423
QY 2413 CTTCCTGGTGTAGTCCACTCTCTGCTATGCTCAATCCCTTCTTACATCTTGTTCAT 2472
Db 2424 CTGCTGGTGTGCTGGCCCTGCTGCTGCTGCTCAACCCACTGCTGTACTGCTCTTCAAC 2483
QY 2473 CCTCACTTTAAGGAGATCTGGTGAGCCT 2501
Db 2484 CCCCACTTCCGGGATGACCTTCGGGGCT 2512

RESULT 12

LOCUS CF742674
DEFINITION UI-M-HB0-clm-q-03-0-UI-r1 NIH_BMAP_HB0 Mus musculus cDNA clone
IMAGE:30617906 5', mRNA sequence.

ACCESSION CF742674.1 GI:37639013

VERSION EST.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..763

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30617906"

/tissue_type="whole eye"

/dev_stage="embryo 12.5,13.5,14.5 dpc"

/lab_host="PH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_HB0"

/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonafide, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is TTATTGAAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 22.2%; Score 604.4; DB 7; Length 763;

Best Local Similarity 87.9%; Pred. No. 7.2e-164;

Matches 670; Conservative 0; Mismatches 91; Indels 1; Gaps 1;

QY 252 TCCTCCCTGCCAGTCTCCGCTTCCTCGAGGAGTTAGCTTTGGGGAACCGCTCTGACATA 311

Db 2 TCTCTCATATCGCTCTGCTTCTTAGAGAGTTAGCTTTGCTGGAATGCTTTGACACA 61

QY 312 CATTCCTCAAGGAGAGATTCATCTGGCTTTTACAGTCTTAAAGTTCTTATCTGCGAGATAA 371

Db 62 CATTCCTCAAGGAGGCTTCAGGGCTTCACAGCTCAAGTGTCTATGCTGAGAACAA 121

QY 372 TCAGCTAAGACAGTACCACAGAGAGCTCTGAGAAATTTGCGAGGCTTCAATCCCTCG 431

Db 122 CCAGCTGAGACAGGTTCCGGAGGAGCGCTACAGAAATTTGAGAGGCTTCAATCCCTCG 181

QY 432 TCTGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTTCAGTGGCTGCTATTCCTC 491

Db 182 CCTAGTGCACACCATCAGCTAGTGCACCCAGCTGTTTCAGGGGCTGCACTCCCT 241

QY 492 GAGGCACTGTGGCTGGATGACAATGCGTTTAAAGAAATCCCGCTCCAGGCTTTTGAAG 551

Db 242 GAGGCACTGTGGCTTAGATGACAATGCTCTACAGAGCTCCCTGTCCAGGCTTTTGAAG 301

QY 552 TTTATCGGATGCAAGCATGACCTTGGCCCTGAAACAAATACACCATACAGACTA 611

Db 302 TTTATCAGCCCTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCATACAGACTA 361

QY 612 TGCCCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACAATAGAAATCCACTC 671

Db 362 CGCCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACAATAGAAATCCACTC 421

QY 672 CTTGGGAAGAAATCTTTGATGGCTCCACAGCCCTAGAGACTTTAGATTTAAATACAA 731

Db 422 CTTGGGAAGAAATCTTTGATGGCTCCACAGCCCTAGAGACTTTAGATTTAAATACAA 481

QY 732 TAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAAGAACTAGGATT 791

Db 482 TAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAAGAACTAGGATT 541

QY 792 TCATAGCAACATATACAGTCTGATACCTGAGAAAGCATTTGTAGCAACCCCTTCTCTTAT 851

Db 542 CCACAGCAACATACAGTCTGATACCTGAGAAAGCATTTGTAGCAACCCCTTCTCTTAT 601

QY 852 TACATACAATTTTATGACAATCCCAATTTGTTGGAGATCTGCTTTTCAACATTT 911

Db 602 CACATACAATTTTATGACAATCCCAATTTGTTGGAGATCTGCTTTTCAACATTT 661

QY 912 ACCTGAACATAGAACACTGACTCTGAAATGGTCCCTCACAAATACTGAATTTCTGATT 971

Db 662 GCCTGAACATAGAACACTGACTCTGAAATGGTCCCTCGCACATTAAGTTTCTCACTT 721

QY 972 AACTGGAATGCAACCTTGGAGACTCTGACTTTTAACTGGAG 1012

Db 722 GACAGGAATGCCACCTTGGAGAGTCTGACTTTTAACTGGAG 763

RESULT 13

LOCUS BX403219

DEFINITION BX403219 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS01029YG09 5-PRIME, mRNA sequence.

ACCESSION BX403219

VERSION BX403219.2 GI:46875014

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tgami, M., Tagawa, A., Takashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Taya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>.

FEATURES source

```

1..3035
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="FANTOM_DB:A930009A08"
    /db_xref="taxon:10090"
    /clone="A930009A08"
    /tissue_type="retina"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="adult"
<1..2638
    /note="unnamed protein product: G PROTEIN-COUPLED RECEPTOR
    LGR4 homolog [Rattus norvegicus] (SPTR|Q922H4, evidence:
    FASTY, 97.2%ID, 92.3%length, match=2634)
    putative"
    /codon_start=2
    /protein_id="BAC31882.1"
    /db_xref="GI:26336394"
    /translation="EDAFKNFDFEELQAGNDLSFHPKALSGIKELKVLTONQL
    KTVFSELRGSLQSLNDLHNTSVPEDEGLVQLRHLWLDNLTVEVPLNSL
    LPTQLALTNLNTSIPDFATNLSSLVHLHNNKIKSLQSHCFGLDLETLDL
    YNNLDEFPQALPSLKEGLFNSISVIPDPAFAGNPLRLTHLVNPLSFVNSA
    FHNLSLHSLVIRGASLVQWPNLGTVHLESLITGTGKISSIPDDLCQOKMLRTLD
    LSYNDIRLPSFNGCRALLESILQBNQISLKEITFFGLTSLRILDLRNLIHHS
    AFKLTITLNDVSNELTSPTEGLNGQLKLVGNFQDLAALDAARDFANRLSVLP
    YAYQCAPFCWCDSYANLNTEDNSPDHSVTKEGADNATSTAEESSQIIHCT
    PSTGAFKPCYGLLGSWMRLTWFIPLVALLNLVILTVFASCSLSPASKFLGLIS
    VSNLMGITYTGILFPLDASVNGRPAEFGIMWETGSKVGLAVFSESASFLLTLA
    AVERSVFAKDYKMGKSHLRFQVAAVALLGAALAGCPFLFHGGQYASAPLCIPLP
    TGETPSLGTVLVLLNSLAEPLLMALITKLYFNLEKEDPSNSOSSMKHVMILFT
    NCIFCPVAFPSFLITLISISPEIMKSVTLIFPLPACLNPLVIFNFPKDDMK
    LLKRVTKHGSVSVSSISSQCGEQDFYDCGMYSHLQGNLTVCDCESFLITKPV
    CKHLKSHSCFLVTASQCPPEAYNSDCGTQSAHSDVADEEDSFVSDSDQVACGRA
    CFYQSRGPFVRYAYNLPVRD"

```

CDS

ORIGIN

```

Query Match      21.2%; Score 576.4; DB 3; Length 3035;
Best Local Similarity 54.5%; Pred. No. 1.8e-155;
Matches 1221; Conservative 0; Mismatches 991; Indels 30; Gaps 2;

Qy 262 AGTCCTCCCTTCCTGGAGAGTTAGCTCTTGGGGAAGCGCTGACATACATATCCCAAG 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17 AATTTCTCTTTCTAGAGAGAGTCACTAGCTGGCTGTGTAACGACCTTCTTTTATCCACCA 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 322 GGAGCATCTCACTGGCCTTTACAGTCTTAAAGTTCTTATGCTGCGACGAATAATCAGCTAAGA 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 AAAGCCTTGTCTGGGTGGAAGAACTCAAGTCTTAACCTCCAGAACAACTAGTTGAAA 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 382 CACGTACCCACAGAGCTCTCGAGAAATTTGCGAAGCCCTTCAATCCCTCGCTCGGATGCT 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

137 ACAGTACCAGTGAAGCCATTCGTGGACTGAGTGTCTTTCAGCTCTCTAGCTTAGATGCC 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 442 AACCAATCATGCTATGTGCTCCCAAGCTGTCTTTCAGTGGCTGCTCATTCCTCGAGGACCTG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 AACCATATTAATCTCAGTCCCGGAGGACAGTCTTGAAGGCTCGTTCAGTTGGGATCTG 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 502 TGGCTGGATGACATGCTGCTTAAAGAAATCCCGCTGCAAGCTTTTAGAAGTTTATCGGCA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 TGGCTGGATGACAAATCTTGAACGAACTGCTGTGCGCTCGCGCTCAGCAACCTGCAACC 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 562 TTGCAAGCCATGACCTGGCCCTGAACAAATACACACATACACAGACTATGCTTGA 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 CTGAGGCGCTGACCTTGGCTCTCAACAAATCTCAAGCATCCCGACTTTCGCTTCAACC 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 622 AACCTCTCCAGCTTGGTAGTCTCATCTCCATCAATAGAAATCCACTCCCTCGGGAAG 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 AACCTTCAAGCTTGGTAGTCTGCATCTTCAACAAATAAATTAAGGCTCAGTCAA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 682 AAATGCTTTGATGGCTCCACAGCTAGAGACTTTTAGATTTTAAATTAACAATACCTTGAT 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 CACTGTTTGTGAGTACTAGATAAACCCTGGAACCCCTGGACTTGAATTAATAAATTTGAT 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 742 GAATTCCTCCACTGCAATTTAGGACACTCTCCAACTTAAAGAACTAGGATTTCAATGCAAC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 GAATTTCTCTCAGGCTATTTAAAGCCCTTCCAGCCCTTAAAGAGCTGGGATTTTCAAGTAA 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 802 AATATCAGGTGATACCTCGAGAAAGACTTTCTGAGCAACCCCTTCTCTTATTAACAATACAT 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 TCTATTTCTGTTATCCCGGATGGAGCATTTGCTGATATCCACTCTAAGAACTATCCAT 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 862 TTCTATGCAATATCCCAATTTGTTGGAGATCTGCTTTTCAACATTTTACCTGAACATA 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 TTGTATGATAATCTCTGCTCTTTTGTGGGAACTCAGCAATTTTCAACACTGTCTGATCTG 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 922 AGAACACTGACTCTGAATGCTGCCTCAAAATTAAGTAACTTCTCTGATTTTAACTGGAAT 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 CATTCCTTTAGTCAATTCGTGGTGCAGAGCCCTGCTGAGTGGTTCCTCCCAATCTGGCCGGA 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 982 GCAAACTCGAGAGTCTGACTTTAACTGGAGCAGAGATCTCATCTCTCTCTCAAAACCGTC 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 GTCCATCTGGAGAGTCTAACTTGCAGGGGACAAAATAAGCAGCATACCTGATGATCTG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1042 TGCAATCAGTTACCTAATCTCAAGTGTAGATCTGTCTTCAACCTATTAGAAGATTTA 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 TGCCAAAACCAAAAGATGCTCGGACTCTGGACTTATCTTATAACGATATAAGAGACCTT 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1102 CCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGACCTAAGACATTAATGAATCTTAC 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 CCAAGTTTTTAATGTTGTCGTCATTTGGAAGAAAATTTTCAITTCGACGCGTAATCAAAATCTCC 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1162 GAAATTTAAAGTTGACACTTTCCAGCAGTGTCTTAGCCCTCCGATCGCTGAATTTGGCTTGG 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 917 CTGATAAAGGAACTACTTTTCAAGGCTAATACCTCCCTAAGGATTTAGATCTGAGTAGA 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1222 AACAAAATGCTATTTATTCACCCCAATGCAATTTTCCATTTTCCCATCCCTAATAAAGCTG 1281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 977 AACCTGATTCGTGAAATTTCAAGTGGAGCTTTTTCGAAAGCTTTGGGACAAATTAACCTG 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1282 GACCTATCGTCCAACTCTCTGCTCTTTTCTTATTAACCTGGGTTACATGGTTTAACTCAC 1341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1037 GATGTGAGTTTCAATGAATTAACCTCAATTTCTTACGGGAAGGCGCTGAATTTGGGCTCAATCAA 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1342 TTAAATTTAAACAGGAAATCATGCTTACAGAGCTTGATATCATCTGAAAAAATTTCCAGAA 1401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1097 CTTAAACTTGTGGGTAACTTCCAGCTGAAAGATGCTTGGCAGCAGCAGAGACTTTGCCAAT 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1402 CTCAGGTTATAGAAATGCTTTATGCTTTACAGTGTCTGTGCTTGGAGTTGGAGTAT 1461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1157 CTCAGGCTCTTATCAGTACCATATGCTTATCAGTGTGTGCAATTTTGGGGGTGTGACTC- 1215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1462 GCCTATAAGATTTCTAATCAATGAATTAAGGTGACAAAGCTATATGGAGCAGCTTCAT 1521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 1216 -----TTATGCAAAATTTAAACACAGAGATAACAGCCCCAGACCAACAGATGTG 1264
QY 1522 AAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAAGTACCTTGAAGATTTCTGCTT 1581
Db 1265 ACAAAGA-----GAAGGTGCTACAGATGCAGCAATGCCACCAGC 1306
QY 1582 GACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAAGTGCAAGTGTTCACCTTCCAGGCC 1641
Db 1307 ACTGCTGAAGTGAAGACATAGCAAAATATCATCATTTGATACACCTTCAACAGGTGCT 1366
QY 1642 TTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATCAGAAATGGAGTGGACCAATA 1701
Db 1367 TTTAAGCCCTGTGAATATTTACTGGGAAGCTGGATGATTCGCCTTACAGTGTGGTTCAAT 1426
QY 1702 CGAGTTCGGACACTTACTTGAATGCTTTGGTGAAGTCTCAACAGTCTTCAGATCCCTCTG 1761
Db 1427 TTCCTGGTGGCTTCTGCTTTTCAACCTGCTTGTCTCATTTTAAAGTGTTCGGTCTTGTTC 1486
QY 1762 TACATTTCCCCCATTTAAACTGTTAATTTGGGTTCATCGCAGCAGTGAACATGCTCACGGGA 1821
Db 1487 TCAGTGCCTGCTCCAGCTCTTCATAGCTTGAATCTGTCTAACTTACTCATGGGC 1546
QY 1822 GTCTCCAGTCCGCTGCTGCTGGTGTGGATGCGTTCACTTTTGGCAGCTTTGCCAGCAT 1881
Db 1547 ATCTATACCTGTCATCTTACTTTTCTGATGCTGTCTCTGGGGCCGATTTGCTGAATTT 1606
QY 1882 GGTCCTGTTGGGAGAAATGGGGTGGTGGCCATGTCATGTTTGTCTCAATTTTGTCT 1941
Db 1607 GGCATTTGGTGGGAACTGGCAGCGGCTGCAAGGTAGCTGGGTCTCTGGCAGTCTTCTCC 1666
QY 1942 TCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGGTCTCTGTGAAA 2001
Db 1667 TCAGAGAGCGGTATTTCTGTAACTCTGCGAGCGGCTGGAAAGAGGTTTTGCAAG 1726
QY 2002 TATTCGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCCTGAAAGTAAATCATTTTGTCT 2061
Db 1727 GATGTAATGAAATTTGGGAAAGCAGTCACCTCGACAGTTCAGGTGGCTGCGCTCGTA 1786
QY 2062 TGTGCGCTGCTGGCTTGCATGACATGGCCGAGTTCCTGCTGGGTGGGAGCAAGTATGGC 2121
Db 1787 GCTTTGCTGGTGTGCAATAGCAGGTGCTTCCCTCTTTCACGGAGGGCAATATTTCT 1846
QY 2122 GCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGAGCCAGCACCATGGGCTACATGGTC 2181
Db 1847 GCATCACCCTTGTGCTGGCTGCTGCTTCTACAGAGAGACACCATATYAGGAATCTACTGTG 1906
QY 2182 GCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATGACCATTTGCCCTACACCAAGCTC 2241
Db 1907 ACCCTAGTGTATTAATCACTAGCAATTTTATGATGGCCATTTATCTACACTAAACTC 1966
QY 2242 TACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGGGACTGCTCTATGGTAAACAC 2301
Db 1967 TACTTCAACTTAGAGAAAGAGACCCGTGAGAAACTCCCACTGTAGCATGATTAAGCAC 2026
QY 2302 ATTGCGCTTGTGCTCTTCAACAACTGCAATCCTAAACTGCTGCTGGCTTCTTGTCCCTTC 2361
Db 2027 GTTGTGCTGCTCATCTTCACAACTGCAATCTTCTGCGCTGTGCAATTTTCTCATTC 2086
QY 2362 TCCTCTTTAAACCTTTACATTTATACATCTCTGAAAGTAAATTAAGTTTATCCCTTCTGCTG 2421
Db 2087 GCACCATTTGATCAGCGCAATCTCCATCAGCCCGAGATAAATGAAGTCTGTTACGCTGATA 2146
QY 2422 GTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCATCTCCTCACTTT 2481
Db 2147 TTTCTTCCGCTGCTGCTGCTGCTGAATCCAGTCTGTAGCTTTCTTCAACCCCAAGTTT 2206
QY 2482 AAGGAGGATCTGGTGAGCTGA 2503
Db 2207 AAAGAGACTGGAAGCTCTCTGA 2228

LOCUS CD617354 590 bp mRNA linear EST 12-JAN-2004
DEFINITION 55041415H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD617354
VERSION CD617354.1 GI:40265619
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 590)
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
AUTHORS Circular rapid amplification of cDNA ends for high-throughput
TITLE extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Pu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source Location/Qualifiers
1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"
ORIGIN
Query Match 21.0%; Score 573.4; DB 6; Length 590;
Best Local Similarity 99.8%; Pred. No. 7e-155;
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 292 GCGGGAACGCTCGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTACAGTCTTAAA 351
Db 16 GCGGGAACGCTCGACATACATTTCCCAAGGGAGCATTTCACTGGCTTTACAGTCTTAAA 75
QY 352 GTTCTTATGCTGCAGAAATAATCAGCTAAGACACACCTACCCACAGAAGCTCTCGAGAATTG 411
Db 76 GTTCTTATGCTGCAGAAATAATCAGCTAAGACACACCTACCCACAGAAGCTCTCGAGAATTG 135
QY 412 CGAAGCTTCAATCCCTCGCTGCTGGATGCTAAACACATCAGCTATGTGCCCTCCAGCTGT 471
Db 136 CGAAGCTTCAATCCCTCGCTGCTGGATGCTAAACACATCAGCTATGTGCCCTCCAGCTGT 195
QY 472 TTCAGTGGCTGCATTCCTCGAGCACCTGCGGCTGGATGACATGCGTTAACAGAAATC 531
Db 196 TTCAGTGGCTGCATTCCTCGAGCACCTGCGGCTGGATGACATGCGTTAACAGAAATC 255
QY 532 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAACAAA 591
Db 256 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAACAAA 315
QY 592 ATACACACATACACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTC 651
Db 316 ATACACACATACACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTC 375
QY 652 CATTAACATAAGATCCACTCCCTCGGAAAGAAATGCTTTGATGGCTCCACAGCTAGAG 711
Db 376 CATTAACATAAGATCCACTCCCTCGGAAAGAAATGCTTTGATGGCTCCACAGCTAGAG 435
QY 712 ACTTTAGATTTAAATTTACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCC 771
Db 436 ACTTTAGATTTAAATTTACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCC 495
QY 772 AACCTTAAAGAACTAGAGATTTATAGCAAAATATCAGGTGATACCTTGAGAAAGCAATTT 831
Db 496 AACCTTAAAGAACTAGAGATTTATAGCAAAATATCAGGTGATACCTTGAGAAAGCAATCT 555
QY 832 GTAGGCAACCTCTCTCTTATTACAAATACATTTCTA 866
Db 556 GTAGGCAACCTCTCTCTTATTACAAATACATTTCTA 590

Search completed: July 12, 2005, 20:26:50
Job time : 9772.8 secs

This Page Blank (uspto)